2004

Thermodynamic characterization of DNA binding by type I DNA polymerases from Thermus aquaticus and Escherichia coli

Kausiki Datta
Louisiana State University and Agricultural and Mechanical College, kdatta1@lsu.edu

THERMODYNAMIC CHARACTERIZATION OF DNA BINDING BY TYPE I DNA POLYMERASES FROM *THERMUS AQUATICUS* AND *ESCHERICHIA COLI*

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Biological Sciences

by
Kausiki Datta
B.S., University of Calcutta, 1995
M.S., University of Calcutta, 1997
May 2004

ACKNOWLEDGEMENTS

I would like to recognize some of the people without whose support and guidance none of this work would have been possible.

Firstly I would like to thank my supervisor, Dr. Vince J. LiCata, for his advice and guidance with my research and the revisions of my writings. Completion of this research and writing of this dissertation would not have been possible without his support and guidance.

I would like to thank all the members of the LiCata lab (both past and present) for being nice and wonderful colleagues.

I would also like to acknowledge all the members of my graduate committee for their helpful discussions and I appreciate the time that each has given to help me complete this work.

Finally I would like to thank my parents and my husband who have supported me in this endeavor.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Taq	*Thermus aquaticus* DNA polymerase I
Klentaq	"Large fragment" of Taq polymerase
Pol I	*E.coli* DNA polymerase I
Klenow	"Large fragment" of *E.coli* Pol I
K_d	Equilibrium dissociation constant
ΔG	Free energy change of binding
ΔH or ΔH_{obs}	Enthalpy change of binding
ΔH_{vH}	van't Hoff enthalpy change of binding
ΔH_{cal}	Calorimetric enthalpy change of binding
ΔS or ΔS_{obs}	Entropy change of binding
ΔS_{vH}	van't Hoff entropy change of binding
ΔCp or ΔCp_{obs}	Heat capacity change of binding
ΔCp_{vH}	van't Hoff heat capacity change of binding
ΔCp_{cal}	Calorimetric heat capacity change of binding
ΔASA	Change in accessible surface area
ΔASA_{pol}	Change in accessible polar surface area
ΔASA_{npol}	Change in accessible non-polar surface area
D	Asparate
E	Glutamate
A	Alanine
[salt]	salt concentration

SK_{obs}	Change in linked ions upon binding
ΔnM^+	Change in the number of monovalent cations
ΔnX^-	Change in the number of monovalent anions
TBP	TATA-box Binding Protein
SSB	Single-strand Binding Protein
Sso7d	A 7 kDa DNA-binding protein from *Sulfolobus solfataricus*
Sac7d	A 7 kDa DNA-binding protein from *Sulfolobus acidocaldarius*
IHF	Integration Host Factor
ORF	Open Reading Frame
CD	Circular Dichroism
SAXS	Small Angle X-ray Scattering
Rg	Radius of gyration

ABSTRACT

DNA binding of the DNA polymerase I from *Thermus aquaticus* (Taq and Klentaq fragment) and *E. coli* (Klenow fragment) have been studied as a function of [salt] and temperature in order to understand their DNA binding thermodynamics. Binding of the two different species of polymerases occurs with sub-micromolar affinities in very different salt concentration ranges. Thus, at similar [KCl] the binding of Klenow is ~ 3kcal/mol (150X) tighter than the binding of Taq/Klentaq. Linkage analysis of the [KCl] dependence of DNA binding at 25^{o}C reveals a net release of 2-3 ions for Taq/Klentaq and 4-5 ions for Klenow. DNA binding of Taq at 60^{o}C only slightly decreases the linked ion release. Linkage analysis of [$MgCl_2$] dependence of DNA binding indicates that formation of protein-DNA complex in both the polymerases is linked to the release of approximately one Mg^{+2} ion.

The temperature dependencies of DNA binding were studied at 5-70^{o}C for Taq/Klentaq, and 5-37^{o}C for Klenow. The temperature dependencies of ΔG of binding by Taq/Klentaq and Klenow show strong curvature due to the presence of negative ΔCp of binding, which was confirmed using isothermal titration calorimetry. DNA binding by both species of polymerase show enthalpy-entropy compensation, with binding being enthalpy driven at their respective physiological temperatures. It is notable that Taq/Klentaq binds DNA as low as 5^{o}C, even though it has almost no catalytic activity at room temperature. Circular dichroism and small angle x-ray scattering measurements show small observable conformational rearrangements upon complex formation for both polymerases. Large negative ΔCp are typically associated with sequence specific DNA

binding. Sequence specificity for the single-stranded template overhang was examined by altering the sequence to poly-A, poly-T and poly-C respectively. The affinity difference from weakest to tightest is only about one order of magnitude. Preliminary studies of the binding of Klenow and Klentaq to different DNA structures show that the two polymerases have different DNA structure preferences. We propose that the relatively large negative ΔCp for Klenow and Klentaq DNA binding might be due to DNA structure specificity or a general characteristic of primarily non-sequence specific DNA binding proteins that bind with high affinity.

CHAPTER 1

GENERAL INTRODUCTION

The faithful replication of the genome is the central element in the maintenance and transmission of genetic information in any organism. This function is accomplished by a family of proteins called "polymerases", that can replicate the genome (DNA or RNA) with extremely high fidelity. The first DNA polymerase to be discovered was DNA polymerase I, isolated from *Escherichia coli,* by Arthur Kornberg in 1955 (1,2). Since then, numerous other DNA polymerases have been isolated and characterized, yet *E.coli* DNA polymerase I remains the central model used for understanding the general mechanism of DNA replication. On the basis of amino acid sequence comparisons (3,4) as well as crystal structure analysis (5), the DNA polymerases are divided into different families (6). Despite the wide variation in the sequences and specific domain structures, all the DNA polymerases utilize an identical two-metal-ion catalyzed polymerization mechanism that adds deoxy-ribonucleotides (dNTP's) to the growing DNA chain during replication (5,6). Interestingly, all the polymerases whose structures are presently known, also appear to share a common overall architectural feature. They resemble a 'half-open' right hand consisting of "thumb", "palm" and "finger" domains (6). The palm domain contains the conserved active site residues important for catalyzing the polymerization reaction, whereas the thumb and the finger domains make important contacts with the incoming dNTP and DNA substrates during polymerization (6). The work performed in this dissertation involves the bacterial type I DNA polymerases from *E.coli* (Pol I) and *Thermus aquaticus* (Taq). Therefore, in this review, I will concentrate primarily on the bacterial type I DNA polymerases. *E.coli* Pol I is the best-studied type I DNA polymerase

and the findings from *E.coli* Pol I are often extrapolated to other polymerase families including the more complex eukaryotic DNA polymerases. Thus, a basic understanding about the *E.coli* Pol I is a starting point for understanding any other DNA polymerases.

E.coli DNA polymerases range from DNA polymerase I, with multiple enzymatic activities in a single polypeptide chain, to the huge multi-subunit complex of DNA polymerase III holoenzyme (comprised of ~ 20 polypeptides with a molecular mass of ~900kDa) where each subunit performs different functions (7,8). In *E.coli* the major replicating polymerase is DNA polymerase III, which synthesizes the leading strand in one continuous piece by adding the dNTP's to the RNA primer, as well as forming the Okazaki fragments during the lagging strand synthesis (9). DNA polymerase I completes the lagging strand synthesis by removing the RNA primers from the Okazaki fragments and simultaneously incorporating the deoxy-ribonucleotides to fill the gaps (9). The nicks left between the consecutive Okazaki fragments are then sealed by DNA ligase, resulting in the formation of a continuous piece of DNA. Thus, Pol I, Pol III and DNA ligase constitute the major DNA replication machinery in *E.coli* (9). Although DNA polymerase III is the major DNA replicating polymerase in *E.coli,* DNA polymerase I is the most abundant *in vivo* (18). Three other DNA polymerases, Pol II (10-12), IV (13) and V (14-16), that are encoded by SOS genes, induced by DNA damage, have also been isolated from *E.coli* (Reviewed in detail in Ref. 17).

Thermus aquaticus (Taq), a thermophilic eubacterium, was discovered by T.D. Brock and co-workers in 1969 (19). This bacterium can survive at temperatures of ~75^{o}C. Unlike many other thermophilic eubacteria and archaea that are known to thrive in extremely saline and acidic conditions (22), *Thermus aquaticus* is found only in

freshwater conditions (19). The DNA polymerase isolated from this bacterium (also known as Taq polymerase) (20,21), is one of the most important biotechnological reagents in use today due to its widespread use in the polymerase chain reaction (PCR) (23). Based on the amino acids sequence similarity, Taq polymerase has been classified in the same family as *E.coli* Pol I, with 38% sequence identity (24). The three dimensional structures of the polymerase domains of the two enzymes are almost identical (See Figure 1.1) and have 49% sequence identity. Like *E.coli* Pol I, Taq polymerase also constitutes the most abundant DNA polymerase in *Thermus aquaticus* (18). The gene sequence of the DNA polymerase III from *Thermus aquaticus* has been recently determined (25), but this protein awaits further characterization.

Thermophilic organisms can be bacterial or archaeal. The archaea are a separate kingdom from the bacteria, and are more closely related to the eukaryotes. Interestingly, as in the case of Taq polymerase, most thermophilic proteins are very similar to their mesophilic homologues, indicating a common ancestral origin (18). From the studies reported in the literature, it has been pointed out that, with the exception of phylogenetic variations, what differentiates thermophilic and mesophilic enzymes is primarily the temperature ranges in which they are stable and active (29), otherwise their sequences are quite similar (ranging from ~40-85%) (24,26), their three dimensional structures are superimposable (27,28, others listed in 29) and they have similar catalytic mechanisms (26,30,31). A number of laboratories have been exploring the basis of the structural stability of extremophilic proteins (for reviews see 29 and 32). This study focuses primarily on the functional characteristics of the mesophilic *E.coli* DNA polymerase I (optimally active at ~ 37^{o}C) and thermophilic Taq polymerase (optimally active at

~75^{o}C). Herein I have focused on the initial step in DNA replication; that is, the binding of the polymerase to DNA and examined the DNA binding thermodynamics, to understand the functional similarities and the differences between the two polymerases and determine if any of these functional characteristics are correlated with the ability of Taq to function at such dramatically higher temperatures than Pol I.

Structure of *E.coli* DNA Polymerase I

E.coli Pol I is a single polypeptide of 928 amino acids and ~103kDa (33) and comprised of three functional domains: the N-terminal 5'nuclease domain (residues ~1-326), the C-terminal polymerase domain (residues ~520-928), and an intermediate proofreading or 3'-5'nuclease domain (residues ~326-519) (24,27). The three-dimensional structure of the full length *E.coli* Pol I has not yet been solved, but a truncated version of the enzyme comprising the polymerase and the proofreading domains, known as the 'large' fragment of Pol I or the Klenow fragment (34), has been crystallized (27) (Figure 1.1). The polymerase domain of the Klenow fragment has the typical "half-open right hand" shape with the 'thumb', 'fingers' and 'palm' subdomains. The high-resolution co-crystal structure of the Klenow fragment bound to DNA shows the DNA bound in an "editing mode" with the 3'end of the primer terminus bound in the proof-reading domain (35) (See Figure 1.2). The 'palm' subdomain, mainly comprised of β-sheets, contain the catalytic residues while the 'thumb' and the 'finger' subdomains are mostly α-helical and are involved in the binding and proper positioning of the DNA and dNTP substrates during the polymerization reaction (5).

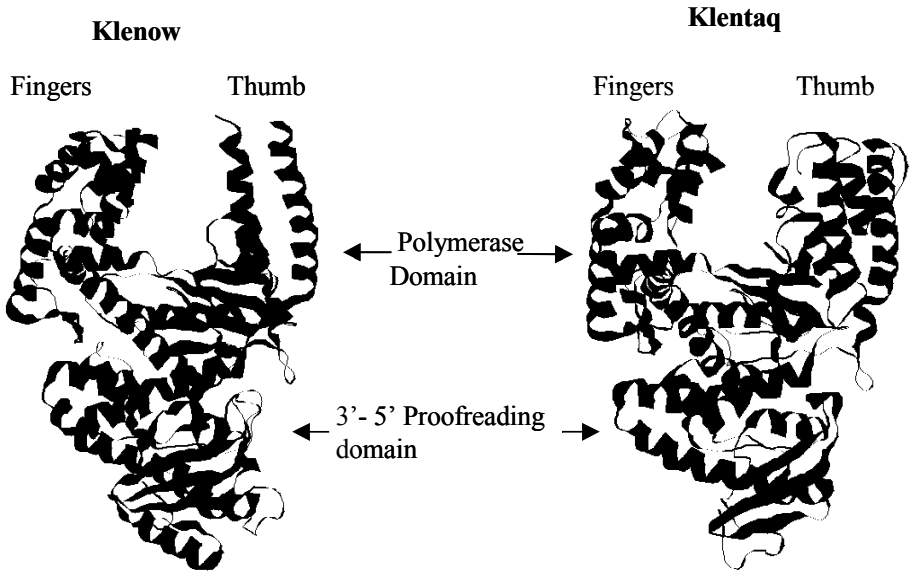


Figure 1.1. X-ray crystal structures of Klenow (Protein Data Bank (PDB) code 1KFD (27)) and Klentaq (PDB code 1KTQ (28)) DNA polymerases.

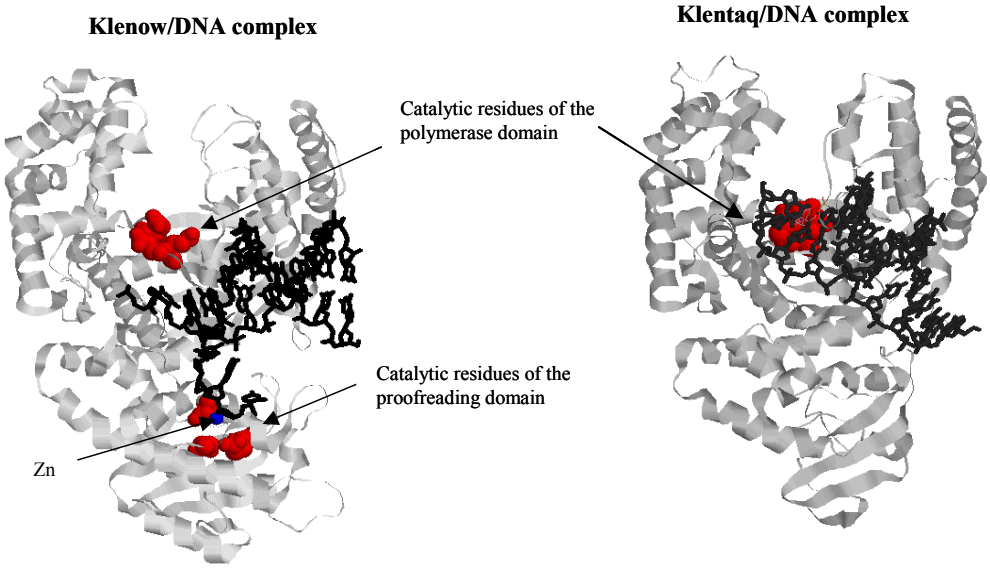


Figure 1.2. X-ray crystal structures of Klenow ((PDB) code 1KLN (35)) and Klentaq (PDB code 4KTQ (51)) polymerases bound to DNA.

Functions of *E.coli* Pol I

Polymerization Reaction and the Proposed Mechanism

All DNA polymerases catalyze the same polymerization reaction where dNTP is added to the growing 3'end of the primer terminus. The addition of the nucleotides is always in the 5'→3' direction and the products are primers with added nucleotides and inorganic phosphate (PPi).

DNA_n+ dNTP = DNA_{n+1} + PPi

The reaction is reversible (pyrophosphorolysis), but the presence of inorganic phosphatases in the cell makes the reaction essentially irreversible *in vivo* (18).

The polymerization reaction involves a nucleophilic attack by the 3'OH of the primer terminus on the dNTP α-phosphate, with the release of pyrophosphate. Benkovic and his group have extensively characterized this reaction pathway for Klenow polymerase and postulated a scheme with seven elementary steps (36). This involves an ordered binding of substrates, with the initial formation of the enzyme-DNA binary complex being required for the productive binding of the dNTP to form a catalytically competent ternary complex. It has also been observed that the actual chemical step (addition of the nucleotide in the growing DNA chain) is both preceded and followed by slow non-chemical processes. These non-chemical steps are believed to be conformational changes in the enzyme or the substrate (or both).

The full kinetic mechanism is summarized as:

$$\mathbf{E + D_n \overset{1}{\Leftrightarrow} E.D_n + dNTP \overset{2}{\Leftrightarrow} E.D_n.dNTP \overset{3}{\Leftrightarrow} E^*.D_n.dNTP \overset{4}{\Leftrightarrow} E^*.D_{n+1}.PP_i \overset{5\,\&6}{\Leftrightarrow} E.D_{n+1} \overset{7}{\Leftrightarrow} E + D_{n+1}}$$

where, E= DNA polymerase, D_n = primer-template DNA, dNTP= deoxy-ribonucleotide, PPi = pyrophosphate, D_{n+1}= extended primer template DNA;

and the steps are:

1. Primer-template binding to the DNA polymerase forming the binary complex,
2. dNTP binding to the binary complex forming the 'open' ternary complex,
3. A conformational change in the enzyme forming the 'closed' ternary complex,
4. The chemical step where the nucleotide is added to the 3'end of the primer strand,
5. A second conformational change of the enzyme,
6. Pyrophosphate release,
7. Either release or sliding of the primer-template for a subsequent round of polymerization

The available structural data combined with mutagenesis studies suggest a polymerase active site structure containing the three crucial carboxylate side chains Asp 705, Asp 882 and Glu 883 in Klenow. It is believed that the carboxylate side chains serve to anchor a pair of divalent metal ions that promote the catalysis. In this mechanism, one Mg^{2+} (number 1) would promote the deprotonation of the 3'OH of the primer strand. The other Mg^{2+} (number 2) would facilitate the formation of the pentacovalent transition state at the α-phosphate of the dNTP and the leaving of the pyrophosphate (5) (See Figure 1.3). Consistent with the idea of catalysis mediated by metal ions, crystallographic experiments have shown the binding of Mg^{2+} and Mn^{2+} to Asp 705 and Asp 882 of Klenow fragment (35,37).

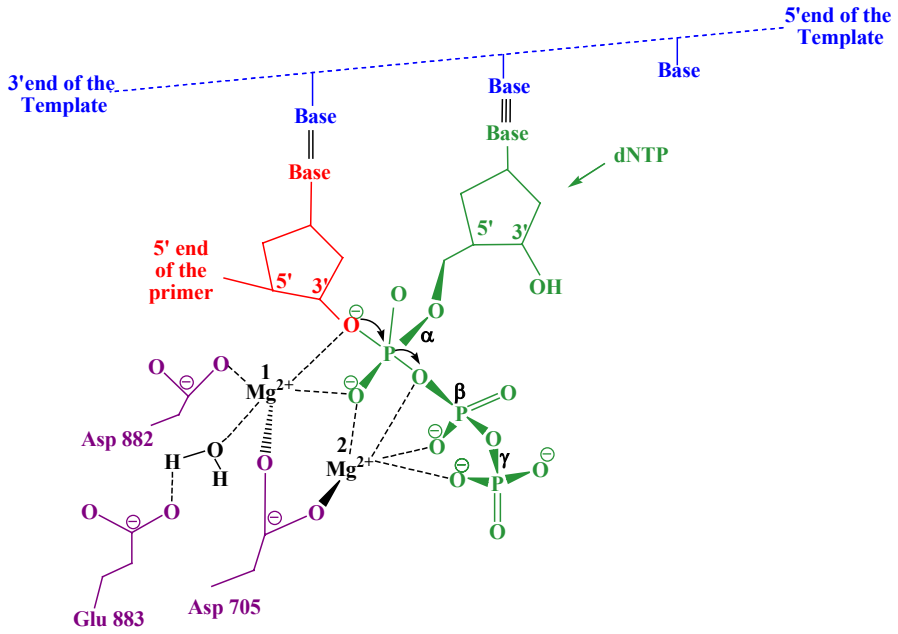


Figure 1.3. Intermediate (or transition state) of the two divalent metal ion mechanism for the polymerase reaction. The roles proposed for Mg^{2+} number 1 are to lower the pK_a of the 3'-OH in order to form the 3'-O^- and to stabilize the 90^o 3'-O-P-O bond angle between the apical and equatorial oxygen atoms. The role of Mg^{2+} number 2 is likewise to stabilize the pentacovalent transition state geometry and to facilitate the leaving of the pyrophosphate. The figure is based on a figure from Steitz, et al., 1994, *Science,* vol. 266, 2022-2025 and is created using ChemDraw.

Proofreading and 5'Nuclease Functions of *E.coli* DNA Polymerase

E.coli Pol I synthesizes DNA with extraordinary fidelity, with an error rate of ~ 1 in 10^9 bases incorporated (38). This high level of accuracy in the replication of DNA is achieved by the 5'nuclease and the proofreading activities that are present in the different subdomains of the polymerase.

The 3'-5' exonuclease component of the polymerase machinery acts in opposition to the direction of DNA synthesis and serves to remove, or proofread, polymerase errors (5,9). The preferred substrate for the exonuclease domain is single-stranded DNA (39) and crystallographic studies have demonstrated the molecular basis for this preference (35,40,41). Thus, the presence of a mismatched nucleotide at the 3'end of the primer

terminus promotes the partitioning of the single-stranded end of the primer terminus to the exonuclease domain of the polymerase, where the mismatch is removed by hydrolysis of the phosphodiester bond (See Figure 1.5).

The 5'-3' exonuclease activity allows the DNA polymerase to extend the 3'terminus of the primer by excising any mismatch that lies in the path of the advancing polymerase (18). *In vivo*, the 5'nuclease activity removes the RNA primers from the Okazaki fragments, during the lagging strand synthesis, by "nick-translation", where the 5'nuclease activity of the polymerase removes the RNA primers and the polymerase fills the gap by incorporating deoxy-ribonucleotides.

Structure and Sequence Comparison Between *E.coli* and Taq DNA Polymerases

The three dimensional structure of Taq DNA polymerase was solved by Steitz and co-workers in 1995 (42). It is classified as DNA polymerase I due to its sequence homology to *E.coli* Pol I and the presence of well conserved active site catalytic residues (5). This along with the DNA bound structure of Taq reported the following year, are the only available crystal structures showing the architecture of a full-length type I eubacterial DNA polymerase. The full-length Taq polymerase is comprised of 832 amino acids with an N-terminal 5'nuclease domain (residues ~1-291), an intervening non-functional proofreading domain (residues ~292-423) and a C-terminal polymerase domain (residues ~424-832), in a single polypeptide chain (See Figure 1.4) (43). The structures of the 'large fragments' (comprising the polymerase and the proofreading domains) of *E.coli* (Klenow) and Taq (Klentaq) (44,45) are almost identical with 49% sequence identity. Although Taq polymerase does not possess any proofreading activity

(20,46), the three-dimensional structure of the proofreading domain is very similar to the proofreading domain of *E.coli* Pol I.

Most of the sequence similarities between the two polymerases are observed in the polymerase and the 5'nuclease domains. Apparently, due to many mutations, deletions, insertions etc. during molecular evolution, the residues corresponding to the proofreading domain of Taq polymerase and *E.coli* Pol I show very little sequence similarity. Taq polymerase is 96 residues shorter than *E.coli* Pol I and most of these inserted residues in *E.coli* occur in the region encompassing the proofreading domain. The structure of Taq polymerase as compared with that of Klenow shows the deletions of four loops of lengths between 8 to 27 residues (42). Out of the three conserved carboxylate residues (Asp 355, Glu 357 and Asp 424), involved in the binding of the divalent metal ions and dNMP during the catalysis of the proofreading activity in *E.coli* Pol I, only Asp 424 has an exact homologue in Taq polymerase (24), although, a different sequence alignment shows that all these metal binding residues are replaced by hydrophobic residues incapable of binding metal ions in Taq (42,45). It is now clear that the DNA polymerases, that have an editing function, possess three small sequence motifs, named Exo I, II and III (47). In Taq DNA polymerase these motifs are completely absent resulting in a non-proofreading polymerase (48). The absence of proofreading activity significantly decreases the fidelity of DNA replication by Taq DNA polymerase, resulting in an error-rate of about 1 misincorporation in 10^5-10^6 bases incorporated (49). The non-proofreading DNA polymerase I in *Thermus aquaticus* (which is evolutionarily separated from *E.coli* by almost 1 billion years) (50) might have allowed the key mutations that led to the evolution of a proofreading *E.coli* Pol I.

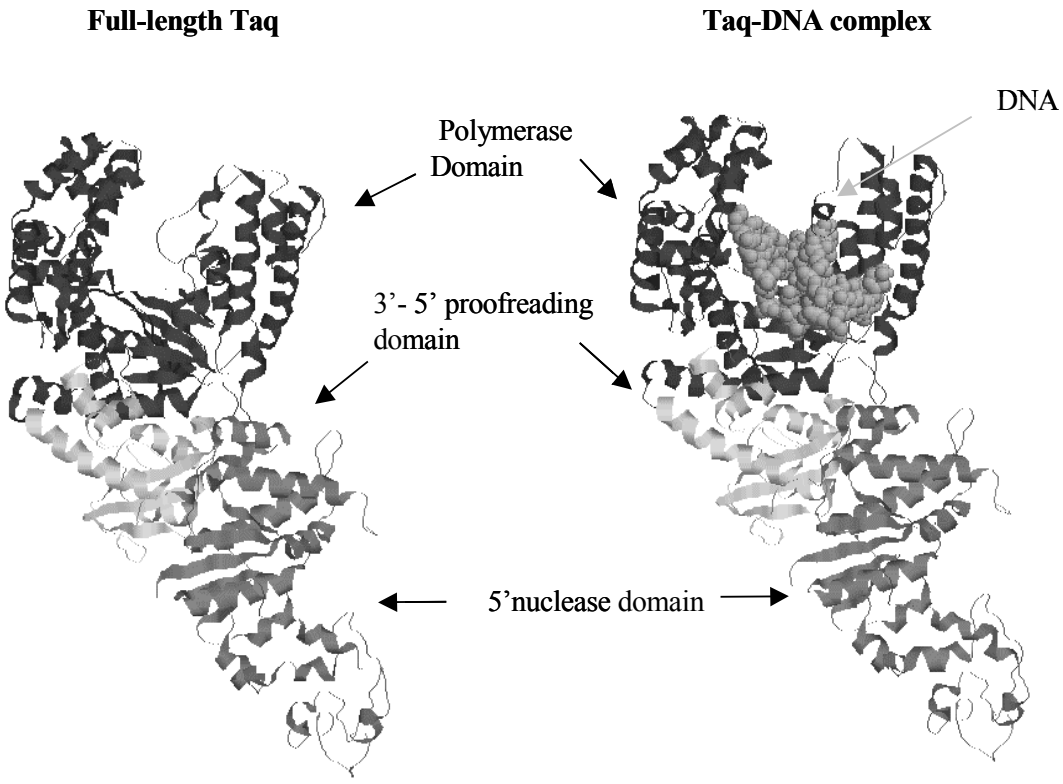


Figure 1.4. X-ray crystal structures of full-length Taq ((PDB) code 1TAQ (42)) and Taq bound to DNA (PDB code 1TAU (62)). In both the structures, the proteins are represented in ribbons with the polymerase domain represented in black, the 3'-5' proof-reading domain in light gray and the 5'nuclease domain in dark gray. The DNA is represented in space-fill form in the structure of Taq-DNA complex.

DNA Binding by Klenow and Taq DNA Polymerase I: A Structural and Biochemical Perspective

Our present understanding of the reaction mechanisms of type I DNA polymerases come from the co-crystal structures of the Klenow (35) and Klentaq (51) with DNA and the in-depth biochemical studies of Klenow DNA polymerase (Reviewed in Ref.5). Although the three-dimensional structures of Klenow and Klentaq DNA polymerases are almost identical, the co-crystals of these polymerases with primed-

template DNA show very different DNA binding modes (35,51) (See Figure 1.2). The structure of the Klenow fragment bound to a primed-template DNA shows the polymerase in an unexpected 'editing mode' where the 3'end of the primer terminus is bound to the proofreading domain and the duplex part of the DNA interacts with a second cleft that runs roughly between the proofreading and polymerase active sites and almost perpendicular to the polymerase domain cleft (35). The structure shows four single-stranded nucleotides at the 3'end of the primer strand bound to the exonuclease active site, identically to the previously described complexes with single-stranded DNA (40,41). The second cleft where the duplex DNA binds is formed partly by significant movement of the thumb subdomain and by reorganization of the region at the tip of the thumb that is highly disordered in the apo structure (52). Extensive interactions are seen across the minor groove between the backbone phosphates of the duplex DNA and amino acid side chains from the thumb subdomain that are highly conserved in the Pol I family (4,5). The conservation of the residues that interact with the duplex DNA implies that the binding site observed in this editing complex is biochemically relevant, and supports the contention that it may also be part of the duplex DNA-binding site employed when DNA is bound in the polymerase mode. In this structure, the primer strand seems to approach the polymerase active site from the direction of the 3' exonuclease domain and the large cleft in the polymerase domain would most probably bind the single-stranded template strand beyond the site of DNA synthesis (5) (discussed in Chapter 5). From the locations of the polymerase active site and the position of the duplex DNA observed in the co-crystal structure, it was predicted that during polymerization the duplex DNA has to bend about 80^{o} in order for the primer terminus to enter the polymerase cleft. Along with these

structural data, the position of the DNA in polymerization mode was predicted from several biochemical and solution based assays. Chemical footprinting (53,54), fluorescence (55,56) and photo-crosslinking (57) experiments together indicate that about 5-8 base pairs of duplex DNA are covered by the Klenow fragment when the primer terminus is at the polymerase active site (5).

In Klenow fragment the polymerase and 3'exonuclease active sites are located about 30Å apart on separate structural domains (35). Although these two catalytic sites usually behave independently of one another (58), the available evidence suggests that the DNA binding functions are shared between the polymerase and the 3'exonuclease domains. Experiments with the separate domains of the Klenow fragment are consistent with the idea of a mutual reliance of one domain on the other for the DNA binding functions (5). Thus, the removal of the 3'exonuclease domain results in a truncated construct with decreased polymerase activity, which is suggested to be due to weakened DNA binding affinity of this truncated polymerase (59,60), while the removal of the polymerase domain leaves an 3'exonuclease domain with no detectable enzymatic or single-stranded DNA binding activity (61). Thus, the observed shift in the duplex DNA binding site together with the interdependence of one domain on the other for DNA binding is consistent with the two site shuttling model proposed for movement of the DNA between the polymerase and the proofreading active sites without dissociation (35) (also discussed later).

The first co-crystal structure of a full length type I DNA polymerase with the DNA bound in the polymerase mode, was the full-length Taq DNA polymerase with a short 8 bp duplex DNA bound in the polymerase cleft (62) (See Figure 1.4). As predicted

from the editing complex of Klenow (35), this structure confirmed that the primer-template DNA approaches the polymerase domain from the 3'exonuclease side of the cleft (62). The duplex DNA shows characteristics of both A and B-form DNA with a more widened minor groove than regular B-DNA. The 3'-OH group of the primer terminus is in close proximity to the three conserved carboxylate residues (Asp 785, Glu 786 and Asp 610) in the polymerase active site. Most of the interactions of the DNA with the polymerase are with the phosphate backbone across the minor groove. There are some conformational changes observed in the protein upon DNA binding specifically in the thumb subdomain so that the DNA is enclosed on three sides by the protein (62). Similar conformational changes are observed for the Klentaq-DNA complex (51). Comparison of the position of the duplex DNA in the polymerase mode in Taq with that of the editing mode in Klenow show that even though they share a common duplex DNA-binding site, the duplexes are translated relative to each other by several angstroms along their helical axes, resulting in different backbone interactions with the proteins (62). This led Steitz to propose that the duplex DNA binding site appears to be constructed such that the DNA can slide by translocation and is not constrained to move only in a screw motion defined by the DNA backbone (62).

In the Klentaq-DNA structure, about nine bases at the 5'end of the template and five bases at the 3'primer terminus of an 11/16mer primed-template DNA participates in contacts with the protein (51). This is similar to the 5-8 bp duplex binding site observed for DNA binding by Klenow in polymerase mode (57). Also, the amino acid residues involved in the binding of DNA in the polymerase mode in Klentaq (determined from the co-crystal structure) (51,62) and Klenow (determined from site-directed mutagenesis

studies) (63,64) are highly conserved, further suggesting that a similar DNA binding mode is employed during polymerization in the two polymerases. The Klentaq co-crystal structure demonstrates that the polymerase domain binds a correctly base-paired primer terminus. Similar conclusions were reported from biochemical studies with Klenow (65,66).

In contrast to the polymerase site, the 3'exonuclease site of Klenow has been shown to bind single-stranded DNA that requires the melting or unwinding of at least four base pairs at the 3'primer terminus of the primed template DNA (35,40,41 and 65,) Structural studies have shown that the four single-stranded nucleotides are bound to the 3'exonuclease site of the enzyme by hydrophobic interactions and hydrogen bonds that are independent of the base sequence (40). Leu 361 and Phe 473 anchor the two 3'terminal nucleotides by stacking with the bases. His 660, contributed by the polymerase domain, interacts with the third base from the 3' terminus, further supporting the contention of interdomain dependence for DNA binding by this polymerase (5,40,41). The sugar phosphate backbone of the single-stranded DNA also interacts with amino acid side chains and the two catalytically important metal ions (40). None of these residues (Leu 361, Phe 473 and His 660) involved in the binding of the DNA in the 3'exonuclease site are conserved in Taq DNA polymerase. The conformation of the single-stranded DNA in the co-crystal structure (40) is much extended and the bases are 'frayed' rather than stacked. The binding site also appears to be 'designed' to bind single-stranded DNA rather than duplex since it is difficult to model build a complementary DNA strand in the binding pocket without creating numerous close contacts with the protein (40,41).

Mechanism of Proofreading Activity and the Role of Divalent Metal Ions

The mechanism of the proofreading reaction is one of the best understood of the reactions catalyzed by DNA polymerase I, due to the availability of the high resolution structural data for co-crystals of Klenow and dNMP (product of the 3'exonuclease reaction) or the single-stranded DNA (the substrate) bound at the exonuclease active site (40,41,52).

The structural and mutagenesis studies demonstrated that, like in the polymerization reaction, the major catalytic role in the proofreading reaction is also played by a pair of divalent metal ions, A and B separated by 3.9Å (40,48). The crystal structure of wild type Klenow with dNMP shows that metal ion A is coordinated to the protein by the carboxylate groups of Asp 355, Glu 357 and Asp 509 and the 5'phosphate of the dNMP, although divalent metal ions (Mg^{2+}, Zn^{2+} or Mn^{2+}) can bind to this site even in the absence of dNMP (40). A second divalent metal ion (site B) is located between the dNMP phosphate and the carboxylate residue Asp 424 only in the presence of dNMP. It is coordinated by the carboxylate of Asp 355, two of the 5' phosphate oxygens of dNMP and also to Asp 424 via bridging water molecules (40). Although crystallographic studies have indicated that both the metal ion binding sites can be filled by Mg^{2+}, Zn^{2+} or Mn^{2+}, in presence of both Zn^{2+} and Mg^{2+}, site A binds Zn^{2+} while site B binds Mg^{2+} (40). However, the exact metal ions used *in vivo* are unknown.

According to the proposed mechanism, metal ion A facilitates the deprotonation of a water molecule thereby forming the hydroxide ion for nucleophilic attack on the phosphodiester bond to be cleaved. After the nucleophilic displacement takes place, both

the metals presumably facilitate the departure of the 3'OH leaving group and are most likely involved in stabilization of the transition state (40,41,48) (See Figure 1.5).

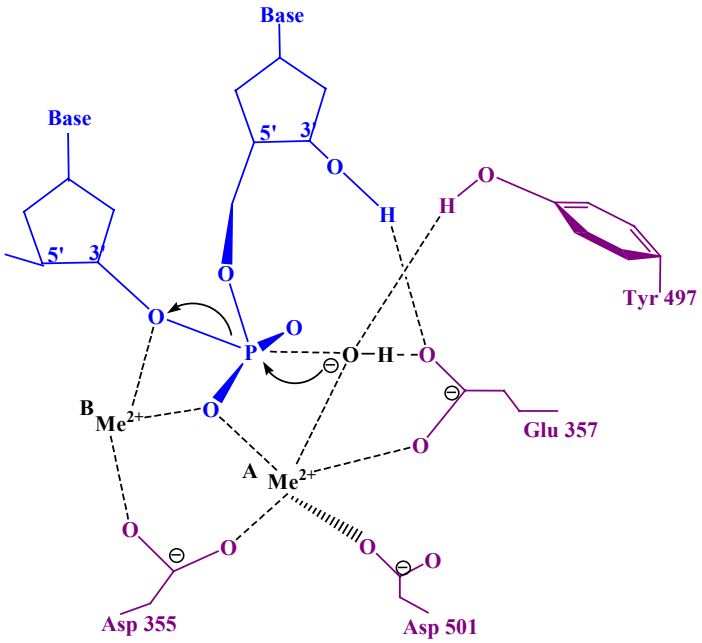


Figure 1.5. The proposed transition state of the two metal ion enzymatic mechanism for the 3'-5'-exonuclease reaction. Metal ion A on the right is proposed to facilitate the formation of an attacking hydroxide ion whose lone pair electrons are oriented towards the phosphorus by interactions with the metal ion, Tyr 497 and Glu 357. Metal B is hypothesized to facilitate the leaving of the 3'-OH group and stabilization of the 90^{o} O-P-O bond angle between the apical and the equatorial oxygen atoms. The figure was created based on a figure in Ref. 40, using ChemDraw.

Mutations of Carboxylate Residues in the Proofreading Domain Results in Klenow Derivatives with Reduced Proofreading Activity

Mutations in the carboxylate residues in the proofreading active site of Klenow fragment resulted in the Klenow derivatives lacking a proofreading function. Crystallographic studies showed that these mutants fail to bind one or more of the divalent metal ions, thus supporting the functional role of these metals in the catalysis.

These mutants are: the D424A single mutant that binds metal A and DNA substrate (or the product, dNMP) in an apparently normal manner but fails to bind metal B (58); the D355A single mutant that binds the DNA substrate and metal A but not metal B (35) and, D355A, E357A double mutant that did not bind dNMP or any of the metal ions (48). These exonuclease deficient mutants are structurally identical to the wild type Klenow fragment, without any effect on the polymerase function (58). Hence, they have proven invaluable in studies involving the polymerase function of the enzyme, where the degradation of the DNA substrate by the exonuclease activity would otherwise interfere with the polymerization reaction.

Present Understanding of the Functional Mechanism of Klenow DNA Polymerase I

The polymerase and the proofreading activities of the Klenow fragment are located on separated structural domains of the protein, separated by about 30Å. An obvious question then is, how the polymerase and the proofreading active sites work together to maintain the fidelity during DNA replication? Although the actual mechanism employed by the enzyme *in vivo* is still unclear, probable models have been proposed, based on structural and biochemical studies.

From structural data, Steitz and Joyce have suggested that the two active sites might be in "communication" by the virtue of the DNA sliding between them without dissociating from the enzyme (41,53). The shuttling of the 3'primer terminus (the substrate for both the polymerase and proofreading activities) from the polymerase to the exonuclease active site, can be achieved by sliding of the DNA over a distance equivalent to 8 base pairs, at the same time allowing 4 base pairs at the primer terminus to become melted (53). Data from biochemical and structural studies have shown unwinding of ~3-4

base pairs at the primer terminus of the DNA is necessary in order to bind to the proofreading domain (41,65). Thus it was proposed that as long as the melting and movement of the primer terminus is fast compared with any of the enzymatic rates (polymerization/exonuclease), the spatial separation of the two active sites need not compromise the efficient coupling of synthesis and editing reactions (41,53).

In addition to the intramolecular shuttling model discussed above, biochemical studies using a trap DNA to limit observations to a single enzyme-DNA encounter (67), have demonstrated that, an intermolecular pathway involving enzyme-DNA dissociation can also be used by Klenow DNA polymerase. This study by Catherine Joyce suggested that the dissociation of the Klenow-DNA complex is the most likely occurrence following a polymerase misincorporation, and editing can be achieved by re-association with the same or a different enzyme molecule, since the "memory" of the DNA synthesis error remains with the product (as a mismatched terminus), not with the enzyme molecule that made the error (67).

A third possibility is that the same enzyme is capable of performing both the polymerase and proofreading functions on different DNA fragments simultaneously. Supporting evidence for this has been provided by Catalano and co-workers, where they have shown that under conditions where the polymerase active site is occupied by duplex DNA, the proofreading domain of the Klenow fragment can bind a second single-stranded DNA substrate and carry out exonucleolytic cleavage (57).

Understanding the Polymerase-DNA Interaction: A Thermodynamic Approach

As discussed above, the functioning of the Klenow DNA polymerase has been extensively investigated at both structural and biochemical levels, but a detailed study at

the thermodynamic level has not been performed. This dissertation presents the first in-depth characterization of the energetics of functioning of this polymerase. In addition, the comparative study of a mesophilic (Klenow) and a thermophilic (Taq/Klentaq) DNA polymerase, performed in this dissertation, has elucidated a number of thermodynamic similarities and differences for polymerases that function in widely different environmental conditions.

In order to understand the molecular basis of any protein-DNA interaction it is essential to understand the non-covalent driving forces that determine the stability and specificity of the protein-DNA complex. Since the energetics of these non-covalent interactions are determined not only by the sequences and the functional groups in the macromolecules but also by the physical environment (i.e. temperature, pH, salt concentration, etc), it is impossible to determine these stabilizing forces based solely on crystal structures. Thus, thermodynamic and kinetic studies as a function of solution conditions are required to be performed to be able to quantify the driving forces that lead to a stable protein-DNA complex. Studies of the thermodynamics of protein-DNA interaction can in principle be examined by two approaches: calorimetry and the measurement of equilibrium binding constants as a function of solution variables (68). Even though calorimetry is the technique of choice for thermodynamicists, as it provides model independent measurements of the energetics of the protein-DNA interaction, for a majority of protein-DNA systems it cannot be successfully used for determining the equilibrium binding constants (typically in the range of 10^6 –10^{12} M^{-1}) with extreme accuracy, primarily because of its requirement for high concentrations (micromolar to millimolar) of reactants. Thus, other solution-based assays, such as nitrocellulose filter-

binding (69), electrophoretic mobility-shift assays (gel-shift) (70) and spectro-fluorometric assays (steady-state fluorescence and fluorescence anisotropy) (71,72), are often employed for obtaining the accurate measurements of the binding constants. Here I have used calorimetric and fluorescence anisotropy assays to characterize the thermodynamics of polymerase-DNA interactions with the type I polymerases from *E.coli* and *Thermus aquaticus*. The use of the two different techniques not only allowed us to obtain additional information, but the independent quantitations also established the validity of the data.

Salt Dependence of DNA Binding Observed for Protein-DNA Interactions

The binding of proteins to nucleic acids is particularly sensitive to the ionic environment of the solution. The highly anionic nature of the DNA results in the local accumulation of cations (mainly K^+ and Mg^{++}) in the vicinity of the DNA. It has been shown that in aqueous media, the fraction of a counterion thermodynamically bound per phosphate is dependent only on the structural charge density along the nucleic acid and the counterion valence and is independent of the bulk salt concentration, as long as it is excess over the phosphate charge (68,73-75). When protein-DNA complexes are formed, the cation-DNA interactions are replaced by protein-DNA contacts with the concomitant release of these cations (73,76), also known as the polyelectrolyte effect. The release of these counterions to the bulk solution is thermodynamically important because it provides a favorable entropic contribution to the binding free energy that stabilizes the protein-DNA complexes (76). With few exceptions (77,78), the equilibrium binding constants (K_{obs}) for protein-DNA associations decrease strongly with the increase in salt concentration [MX]. Plots of $\ln K_{obs}$ versus ln[MX] (also known as linkage plots, 79), are

typically linear at low salt concentrations (usually ≤0.5M), in the absence of any competitive equilibria associated with the binding reaction (80). The dependence of the K_{obs} on salt concentration is given by the following equation, which describes the slope (SK_{obs}) of a linkage plot and is a quantitative estimate of the net number of thermodynamic ions released or taken up when the protein-DNA complex is formed.

$$SK_{obs} = (\partial \ln K_{obs}/\partial \ln[MX]) = \Delta n M^{+} + \Delta n X^{-}$$

Since in majority of the protein-DNA interactions, the binding affinity decreases with increasing [salt], the slope of the linkage plot is typically negative, showing that there is a net release of ions upon complex formation.

Even though SK_{obs} is the measure of the sum of the stoichiometries of both cations and anions released upon binding, it has been shown that the release of the cations is usually the primary determinant of the SK_{obs} for protein-DNA interactions (discussed in details in Ref. 80). This is largely due to the fact that the majority of the ions that are released upon binding, are the cations that shielded the high density of negative charges on the DNA in solution. On the other hand, the domains of the proteins that interact with the DNA are typically relatively short oligocations with much smaller number of charges than that of polyanionic DNA. Thus, cation release from the DNA is typically the major component of SK_{obs}. It is important to note here, that the dependence of the K_{obs} on the salt concentration, observed for protein-DNA interactions, arises from the direct binding or release of specific ions, as discussed above, and not due to a simple ionic strength effect. This is most clear from the fact that dramatically different values of K_{obs} are obtained in buffers containing monovalent cations versus divalent cations, even though the solutions are at same ionic strength (68).

Specific Ion Effects in Protein-DNA Interaction

A) Cation Specific Effects (Monovalent Versus Divalent) on the Binding of Proteins to DNA

For protein-DNA interactions for which the effects of different monovalent cations have been studied, the nature of the cation is shown to have no effect on the K_{obs} or SK_{obs} (73,76,81). This is consistent with the polyelectrolyte theory. However, the lack of a larger effect of the nature of the cation at fixed salt concentration, is somewhat surprising, since ^{23}Na NMR investigations of competitive titrations of NaDNA with univalent salts demonstrated a clear hierarchy of affinities of cations for the vicinity of the DNA. Relative affinities decreased according to the series $NH_4^+ > Cs^+ > Rb^+ > K^+ > Li^+ > Na^+ > Et_4N^+$(tetraethyl ammonium ion) $> Bu_4N^+$ (tetra-n-butyl ammonium ion) (82,83).

The effects on the K_{obs} caused by the divalent cations, such as Mg^{2+}, are qualitatively similar to that of monovalent cations, as discussed above. However, since the number of divalent counterions thermodynamically associated per phosphate of the DNA is approximately one-half that of the monovalent counterions, the SK_{obs} in the presence of Mg^{2+} will be one-half its value in monovalent salt. A deviation from this indicates that the divalent cation plays an additional specific role in complex formation (68). The divalent cations compete more effectively than monovalent cations for accumulation near the surface of the DNA. Thus, in buffers containing a mixture of monovalent and divalent cations (often used for studying protein-DNA interactions), the absolute value of the SK_{obs} for monovalent ions will be lower than in the absence of divalent ions, because when both M^+ and $Mg^{2+}$are bound to the nucleic acid, less M^+ will be released on formation of the complex (68).

In this dissertation I have characterized the KCl and $MgCl_2$ dependencies of DNA binding by Taq/Klentaq and Klenow DNA polymerases using a fluorescence anisotropy based binding assay and have quantitated the linked ion releases upon DNA binding by these polymerase (See Chapter 3). Our results show that in spite of the significant structural similarity, these two species of DNA polymerases, differ in their intrinsic DNA binding affinity, their salt sensitivity, their linked ion releases and their need for bound divalent cations.

B) Effects of Anions (Especially Glutamate) on DNA Binding

For some protein-DNA interactions the nature of anions have been shown to have a large effect on the binding affinity (K_{obs}) with little or no effect on the SK_{obs} (81). The protein-DNA system for which the anion effects have been most extensively studied are the specific and non-specific binding of the lac repressor (lac R) (80). For both specific and non-specific binding of lac R to the wild-type lac operon, replacement of chloride (Cl^-) with acetate (Ac^-), leads to an approximately 30 fold increase in K_{obs} at constant salt concentration with ~10-15% decrease in the SK_{obs} (84,85). At a fixed salt concentration of 0.13M, K_{obs} for the lacR-lacO^+ repressor-operator specific interaction was found to decrease overall by a factor of $\sim 10^4$ in the order $Ac^- \geq F^- > Cl^- > Br^- > NO_3^- > SCN^- > I^-$, as the nature of the anion was varied from the weak-interacting to the strong-interacting end of the Hofmeister series (84), with no significant effect on the $SK_{obs.}$

Historically, most studies of *E.coli* protein-DNA interactions have been performed in NaCl or KCl *in vitro*, as the nature of the anion did not appear to affect the $SK_{obs.}$ But for *E.coli* the physiologically important anion has been proposed to be glutamate (Glu^-) rather than Cl^- (81). This led to the determination of the effect of

replacing Cl^- by Glu^- on values of K_{obs} and SK_{obs} of various protein-DNA interactions. For most proteins examined to date, K_{obs} (from direct measurements or extrapolation) are found to be substantially higher (somewhere between 10X to 800X) in Glu^- than in Cl^- at any specified salt concentration (81,86,87). For some of the enzymes, the high K_{obs} has been correlated to higher enzymatic activity in Glu^- compared to Cl^-, but the molecular basis of the increased affinity and/or increased enzymatic activity in Glu^- compared to other anions is not clear. The replacement of Cl^- with Glu- also lowered the SK_{obs} in some of these reactions, the effect being most prominent for the specific binding of lac R to O^{sym} operator where the SK_{obs} is decreased by 40% is observed (81).

The decreased SK_{obs} in presence of Ac^- or Glu^- compared to Cl^- might be attributed to the differential anion binding by the protein and/or the complex, but no unique interpretation of the large effects on the K_{obs} has yet been possible, primarily because the reference point of a non-interacting anion is not known (80). However, an understanding of these anion effects is essential in order to further understand the contribution of the polyelectrolyte effect towards the stability of the protein-DNA complex.

Temperature Dependence of DNA Binding Observed for Different Protein-DNA Interactions

The stability of a protein-DNA complex at any temperature is defined by the free energy change (ΔG_{obs}) of the association process and can be obtained directly from the equilibrium binding constant (K_{obs}) at that temperature:

$$\Delta G_{obs} = -RT\ln K_{obs}$$

where, R is the universal gas constant and T is the temperature.

However, in order to understand the non-covalent interactions and the nature of thermodynamic driving forces for a particular macromolecular binding reaction, we need to dissect the ΔG_{obs} for the association reaction into the enthalpic (ΔH_{obs}) and entropic (ΔS_{obs}) contributions at the temperature of interest. A careful determination of the temperature dependence of K_{obs} (van't Hoff analysis, discussed below) provides a complete description of these enthalpic and entropic parameters. Although the enthalpy (ΔH_{obs}) of the process can also be measured directly and more accurately by isothermal titration calorimetry, due to the limitations of this assay, van't Hoff analysis is more commonly used for obtaining the energetic parameters of interest for protein-DNA interactions. (80,88)

van't Hoff and/or Gibbs-Helmholtz Analysis of the Temperature Dependence of K_{obs} for Protein-DNA Association

The temperature dependence of K_{obs} is most commonly represented as a van't Hoff plot ($\ln K_{obs}$ versus 1/T) or a Gibbs-Helmholtz plot (ΔG versus T). For macromolecular binding processes studied over a narrow range of temperatures, the van't Hoff and the Gibbs-Helmholtz plots are often linear. The ΔH_{obs} and the ΔS_{obs} can be calculated from the slopes and intercepts of these plots respectively,

$$\ln K_{obs} = -\Delta H_{obs}/RT + \Delta S_{obs}/R \qquad \text{(van't Hoff plot)}$$

$$\Delta G_{obs} = \Delta H_{obs} - T\Delta S_{obs} \qquad \text{(Gibbs-Helmholtz plot)}$$

A linear van't Hoff/ Gibbs-Helmholtz plot shows that both ΔH_{obs} and ΔS_{obs} of the reaction are independent of temperature (at least over the temperature range examined).

In contrary to this simple behavior, many protein-DNA interactions (especially sequence-specific protein-DNA interactions), investigated as a function of temperature,

exhibit a highly non-linear van't Hoff /Gibbs-Helmholtz plot (89-96). This arises due to a strong temperature dependence of the ΔH_{obs} for the binding reaction that is described by the heat capacity change (ΔCp_{obs}) associated with the process, $(\partial \Delta H_{obs}/\partial T) = \Delta Cp_{obs}$. Thus, a non-linear analysis (shown below) of the temperature dependence of either K_{obs} (van't Hoff plot) or ΔG_{obs} (Gibbs-Helmholtz plot) allows us to determine the ΔH_{obs} and ΔS_{obs} of binding as a function of temperature and the ΔCp_{obs} associated with the process (discussed in datail in Chapters 4 and 5 of this dissertation).

Non-linear van't Hoff analysis:

$$\ln(K_{obs}/K_{refT}) = \{(\Delta H_{refT} - T_{refT}\,\Delta Cp)/R\}(1/T_{refT} - 1/T) + \Delta Cp/R\,\ln(T/T_{refT})$$

where, K_{obs} is the binding constant for DNA binding (dependent variable), K_{refT} is the binding constant at any chosen "reference temperature" T_{refT}, T is the temperature in Kelvin (independent variable), ΔCp is the heat capacity, and ΔH_{refT} is the fitted van't Hoff enthalpy at T_{refT}.

Non-linear Gibbs-Helmholtz analysis:

$$\Delta G_{(T)} = \Delta H_{refT} - T\Delta S_{refT} + \Delta Cp\,[T - T_{refT} - T\,\ln(T/T_{refT})]$$

where, $\Delta G_{(T)}$ is the free energy at each temperature (dependent variable), T is the temperature in Kelvin (independent variable), ΔCp is the heat capacity, and ΔH_{refT} and ΔS_{refT} are the fitted van't Hoff enthalpy and entropy values at any chosen "reference temperature" T_{refT}.

Significance of the ΔCp of DNA Binding

One of the challenging goals of thermodynamics is to be able to predict the molecular mechanism of a macromolecular process based primarily on the energetic information. The energetics of the protein-DNA interactions most extensively

characterized to date, are that of the sequence-specific DNA binding proteins (such as restriction endonucleases, transcription repressors and activators, etc). These proteins bind preferentially to specific recognition sites on the DNA with very high affinity and also bind more weakly to non-specific DNA sequences. The energetic parameters for these two modes of DNA binding are significantly different and the basis for these differences at the molecular level were recently reviewed for several systems (97).

Specific binding is characterized by a strong temperature dependence of the binding enthalpy, resulting in a highly negative ΔCp that ranges typically from about –0.36 to –3.4 kcal/mole K for the systems studied (89-96). It had been shown previously that for protein folding reactions, negative ΔCp could be correlated to the burial of non-polar surfaces upon folding (98-103). The crystal structures of the specific protein-DNA interactions are also characterized by a tight complementary recognition interface that buries significant amounts of non-polar surfaces (97). This prompted Record and co-workers (88,104) to suggest that the hydrophobic effect was primarily responsible for the large negative ΔCp observed for proteins binding to specific DNA sites. In contrast, the non-specific mode of binding by these proteins is characterized by $\Delta Cp \approx 0$, a rigid-body association between the protein and the DNA and a much smaller interface area in the complex (105,106). Thus, the strong negative ΔCp is considered to be a 'thermodynamic signature' of sequence-specific DNA binding (97), and has proven to be a powerful connection between thermodynamic and structural information. The structural feature that is most frequently correlated with the negative ΔCp, to establish this connection, is the change in the water accessible surface area (ΔASA) upon complex formation. From studies on transfer of hydrocarbons to aqueous environment and protein

folding/unfolding reactions, quantitative relationships between ΔCp and ΔASA have been proposed (98-103). For many DNA-binding proteins, however, the ΔASA values calculated from the ΔCp using these empirical relationships are much larger than those calculated from the crystal structures (88-91,93,94,107). In other words, the ΔASA calculated for the burial of surface area at the interface, based on crystal structures, could not account for all the observed negative ΔCp. One of the probable origins of this discrepancy might be the highly polar nature of protein-DNA interfaces as opposed to interiors of proteins. Although the empirical relations have been refined to account for the contributions from polar and non-polar surfaces to the total ΔASA for protein folding reactions (100-103), the ratio of these contributions might be significantly different for a protein-DNA interface. Thus, using these correlations for protein-DNA interactions might not be appropriate. Presently, the origin of the observed discrepancy is not well understood, but it does appear that these correlations alone cannot account for all the observed negative ΔCp in the majority of protein-DNA interactions, and that additional linked processes also need to be considered in order to explain the molecular origin of the negative ΔCp. Although all of these processes are not correlated as quantitatively as the ΔCp versus ΔASA correlation, they have been shown to make significant contributions to the observed ΔCp.

Other Probable Molecular Contributions for Negative ΔCp of DNA Binding

A) Conformational Changes in the Protein upon Binding:

As discussed above, the ΔASA due to burial of surfaces at the interface could not account for all the observed negative ΔCp. This led Spolar and Record (104) to suggest the 'coupled binding and folding model', to explain the molecular origin of the high

negative ΔCp. They proposed that binding induced the folding of local regions of the protein (analogous to disordered to ordered transitions in the free and bound enzymes), and that this also caused the burial of large amounts of non-polar surfaces in addition to the interface ΔASA. Although this was successful in accounting for the negative ΔCp in some protein-DNA interactions, it could not explain the discrepancy in a vast majority of the protein-DNA interactions for which both thermodynamic and structural data are available (See Ref. 97 for specific examples).

B) Distortion or Conformational Rearrangement of the DNA upon Binding

Another major proposed origin of the negative ΔCp for protein-DNA interactions is the conformational changes of the DNA upon binding. Binding of TBP to the E4 promoter has been shown to be associated with a large negative ΔCp that could not be explained by significant burial of non-polar surfaces or conformational changes in the protein (93). Here they found that upon binding of TBP, the TATA box DNA is distorted from the regular B form, resulting in a DNA that is unwound, severely bent and doubly kinked, leading to the widening of the minor groove and the creation of the surface that is complementary to the TBP-DNA binding interface (93). Ferrari and Lohman (107) also proposed that the observed negative ΔCp for the non-specific interactions between *E.coli* SSB protein and poly (dA), is partially the result of unstacking of the DNA bases and the binding of the protein to the fully unstacked DNA.

C) Restriction of the Degrees of Freedom of the Complementary Polar, Hydrated Surfaces at the Protein DNA-Interface

Based on the calorimetric results for the trp repressor-DNA binding (89), it was suggested that in addition to the hydrophobic effect, the restriction of degrees of freedom of the interfacial polar side chains, nucleotide bases and solvent exposed backbone

elements, as well as their associated hydration sites, must make a significant contribution to the negative ΔCp of binding. High-resolution structural data for the trp repressor-DNA complex indicate the presence of many trapped water molecules at the interface that interact with the protein, DNA or other water molecules via a network of hydrogen bonds. Many of these water molecules mediate protein-base and protein-phosphate interactions that supplement the direct protein-DNA contacts, thereby dictating the specificity of this interaction. It was hypothesized that the presence of these water molecules provides an additional contribution to the ΔCp through their restriction in the interface (89). This hypothesis was further supported by the data of MetJ repressor-DNA interaction, suggesting that interfacial water molecules contribute significantly to the thermodynamic properties of an interaction (reviewed in detail in Ref. 94).

In this dissertation I have examined the temperature dependence of equilibrium DNA binding by Taq/Klentaq and Klenow DNA polymerases in order to understand the energetics of the reaction (reported in Chapter 4 and 5). These polymerases are known to bind DNA primarily in a non-sequence specific manner and this non-specificity seems to be a pre-requisite for the repair and replication of DNA within the cell. Interestingly, DNA binding by both the polymerases is associated with high negative ΔCp's, that are comparable to sequence-specific binding proteins. The probable molecular origins for the negative ΔCp values are also discussed (See Chapters 4 and 5). Our results also show, that in spite of the wide separation in their physiological temperatures, both Taq and Klenow polymerases bind primer-template DNA with very similar thermodynamic profiles. It is also notable from this study that although Taq DNA polymerase has no

measurable catalytic activity below room temperature (108,109), it can bind DNA quite well down to at least 5 ^{o}C.

Although the thermodynamic profile of the binding of Taq and Klenow DNA polymerases to primed template DNA are very similar, our preliminary studies show that their binding affinities for different DNA structures are very different (See Appendix 1). We find that the Klenow DNA polymerase binds perfect double stranded DNA with extremely weak affinity as compared to primed template DNA. Similar observations were previously reported for double-stranded DNA binding by Klenow (110) and T4 DNA polymerase (111). In contrast, our data shows that Taq DNA polymerase can bind double-stranded DNA with high affinity comparable to that of primed-template DNA. Interestingly, the heat capacity changes for binding of Klenow to the different DNA structures are also different, suggesting different structure-specific binding modes.

References

1. Kornberg, A., Lehman, I.R., Bessman, M.J. and Simms, E.S. (1956) *Biochim. Biophys. Acta.* **21**, 197-198.

2. Lehman, I.R., Zimmerman, S.B., Adler, J., Bessman, M.J. and Simms, E.S. (1958) *Proc. Natl. Acad. Sci. U.S.A.* **44**, 1191-1196.

3. Delarue, M., Poch, O., Tordo, N., Moras, D. and Argos, P. (1990) *Protein Eng*. **3**, 461-467.

4. Braithwaite, D.K. and Ito, J. (1993) *Nucleic Acids Res*. **21**, 787-802.

5. Joyce, C.M. and Steitz, T.A. (1994) *Annu. Rev. Biochem*. **63**, 777-822.

6. Steitz, T.A. (1999) *J. Biol. Chem*. **274**, 17395-17398.

7. Kornberg, A., Baker, T.A., Bertsch, L.L., Bramhill, D., Funnell, B.E., Lasken, R.S., Maki, H., Maki, S., Sekimizu, K. and Wahle, E. (1987) in *DNA Replication and Recombination* (McMacken, R., and Kelly, T.J., eds.) pp 137-149, Alan R. Liss, Inc., New York.

8. Maki, H., Maki, S. and Kornberg, A. (1988) *J. Biol. Chem*. **263**, 6570-6578.

9. Kornberg, A. and Baker, T.A. (1992) *DNA Replication*. New York: Freeman. 2nd Ed.

10. Bonner, C.A., Randall, S.K., Rayssiguier, C., Radman, M., Eritja, R., et. al. (1988) *J. Biol. Chem*. **263**, 18946-18952.

11. Bonner, C.A., Hays, S., McEntee, K. and Goodman, M.F. (1990) *Proc. Natl. Acad. Sci. U.S.A*. **87**, 7663-7667.

12. Iwasaki, H., Nakata, A., Walker, G. and Shinagawa, H. (1990) *J. Bacteriol*. **172**, 6268-6273.

13. Wagner, J., Gruz., P., Kim, S.R., Yamada, M., Matsui, K., et. al. (1999) *Mol. Cell*. **40**, 281-286.

14. Tang, M.J., Bruck, I., Eritja, R., Turner, J., Frank, E.G. et. al., (1998) *Proc. Natl. Acad. Sci. U.S.A*. **95**, 9755-9760.

15. Tang, M.J., Shen, X., Frank, E.G., O'Donnell, M., Woodgate, R. and Goodman, M.F. (1999) *Proc. Natl. Acad. Sci. U.S.A*. **96**, 8919-8924.

16. Reuven, N.B., Arad, G., Maor-Shoshani, A. and Livneh, Z. (1999) *J. Biol. Chem*. **274**, 31763-31766.

17. Goodman, M.F. (2002) *Annu. Rev. Biochem*. **71**, 17-50.

18. Perler, F.B., Kumar, S and Kong, H. (1996) *Adv. Protein Chem*. **48**: 377-435.

19. Brock, T.D. and Freeze, H. (1969) *J. Bacteriol*. **98**, 289-297.

20. Chien, A., Edgar, D.B. and Trela J.M. (1976) *J. Bacteriol*. **127**, 1550-1557.

21. Kaledin, A.S., Sliusarenko, A.G. and Gorodetskii, S.I. (1980) *Biokhimia*. **45**, 644-651.

22. Brock, T.D., Brock, K.M., Belly, R.T. and Weiss, R.L. (1972) *Arch, Mikrobiol***. 84,** 54-68.

23. Saiki, R.K., Gelfand, D.H., Stoffel, H., Scharf, S., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) *Science* **239**, 487-491.

24. Lawyer, F.C., Stoffel, S., Saiki, R.K., Myambo, K., Drummond, R. and Gelfand, D.H. (1989) *J. Biol. Chem*. **264**, 6427-6437.

25. Huang, Y.P. and Ito, J.J. (1999*) Mol. Evol*. **48**, 756-769.

26. Vieille, C., Hess, J.M., Kelly, R.M. and Zeikus, J.G. (1995) *Appl. Environ. Microbiol.* **61**, 1867-1875.

27. Beese, L.S., Friedman, J.M. and Steitz, T.A. (1993) *Biochemistry* **32**, 14095-14101.

28. Korolev, S., Nayal, M., Barnes, W.M., Di Cera, E. and Waksman, G. (1995) *Proc. Natl. Acad. Sci.USA*. **92**: 9264-9268.

29. Vieille, C. and Zeikus, J.G. (2001*) Microbiol. Mol. Biol. Rev*. **65**, 1-43.

30. Zwickl, P., Fabry, S., Bogedain, C., Hass, A. and Hensel, R. (1990) *J. Bacteriol*. **172**, 4329-4338.

31. Bauer, M.W. and Kelly, R.M. (1998) *Biochemistry* **37**, 17170-17178.

32. Sterner, R. and Liebl, W. (2001) *Crit. Rev. Biochem. Mol. Biol*. **36**, 39-106.

33. Jovin, T.M., Englund P.T. and Bertsch L.L. (1969) *J. Biol. Chem*. **244**, 2996-3008.

34. Klenow, H. and Henningsen, I. (1970) *Proc. Natl. Acad. Sci*. *U.SA*. **65**, 168-175.

35**.** Beese, L.S., Derbyshire, V. and Steitz, T.A. (1993) *Science* **260**, 352-355.

36. Kuchta, R.D., Mizrahi, V., Benkovic, P.A., Johnson, K.A. and Benkovic, S.J. (1987) *Biochemistry*. **26**, 8410-8417.

37. Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A. and Steitz, T.A. (1992) *Science* **256**, 1783-1790.

38. Echols, H. and Goodman, M.F. (1991) *Annu. Rev. Biochem*. **60**, 477-511.

39. Brutlag, D. and Kornberg, A. (1972) *J. Biol. Chem.* **247**, 241-248.

40. Beese, L.S. and Steitz, T.A. (1991) *EMBO J*. **10**, 25-33.

41. Freemont, P.S., Friedman, J.M., Beese, L.S., Sanderson, M.R. and Steitz, T.A. (1988) *Proc. Natl. Acad. Sci*. *U.S.A*. **85**, 8924-8928.

42. Kim Y., Eom S.H., Wang J., Lee D.S., Suh S.W. and Steitz T.A. (1995) *Nature* **376**, 612-616.

43. Li, Y. and Waksman, G. (2001) *Curr. Org. Chem.* **5**, 871-883.

44. Barnes, W.M. (1992) *Gene (Amst.)* **112**, 29-35.

45. Korolev, S., Nayal, M., Barnes, W.M., Di Cera, E. and Waksman, G. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9264-9268.

46. Lawyer, F.C., Stoffel, S., Saiki, R.K., Chang, S.Y., Landre, P.A., Abramson, R.D. and Gelfand, D.H. (1993) *PCR Methods Appl.* **2**, 275-287.

47. Bernad, A., Blanco, L., Lazaro, J.M., Martin, G. and Salas, M. (1989) *Cell* **59**,219-228.

48. Derbyshire, V., Pinsonneault, J.K. and Joyce, C.M. (1995) *Methods Enzymol*. **262**, 393-385.

49. Eckert, K.A. and Kunkel, T.A. (1990) *Nucleic Acids Res*. **18**, 3739-3744.

50. Battista, J. R. and F. A. Rainey. (2001) in *Bergey's Manual of Systematic Bacteriology* (Boone, D. R., and Castenholz, R.W., eds), 2nd Ed., pp. 395-414, Vol. 1, Springer-Verlag New York Inc., New York.

51. Li, Y., Korolev, S. and Waksman, G. (1998) *EMBO J.* **17,** 7514-7525.

52. Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G. and Steitz, T.A. (1985) *Nature* **313**, 762-766.

53. Joyce, C.M. and Steitz, T.A. (1987) *Trends Biochem. Sci*. **12**, 288-292.

54. Joyce, Ollis, Rush and Steitz, Koningsburg and Grindley, (1986) Protein structure, Function and Design, UCLA Symposium on Molecular and Cellular Biology, (Oxender D Ed) vol **32**, 197-205.

55. Allen, D.J., Darke, P.L. and Benkovic, S.J. (1989) *Biochemistry* **28**, 4601-4607.

56. Guest, C.R., Hochstrasser, R.A., Dupuy, C.G., Allen, D.J., Benkovic, S.J. and Millar, D.P. (1991) *Biochemistry* **30**, 8759-8770.

57. Catalano, C.E., Allen, D.J. and Benkovic, S.J. (1990) *Biochemistry*. **29**, 3612-3621.

58. Derbyshire, V., Freemont, P.S., Sanderson, M.R., Beese, L.S., Friedman, J.M., Joyce, C.M. and Steitz, T.A. (1988) *Science* **240**, 199-201.

59. Freemont, P.S., Ollis, D.L., Steitz, T.A. and Joyce, C.M. (1986) *Proteins* **1**, 66-73.

60. Derbyshire, V., Astatke, M. and Joyce, C. (1993) *Nucleic Acids Res.* **21**, 5439-5448.

61. Derbyshire, V. (1990) Studies of the 3' to 5' exonuclease of DNA polymerase I of E.coli. PhD thesis, Yale University, New Haven. 111pp.

62. Eom, S.H., Wang, J. and Steitz, T.A. (1996) *Nature* **382,** 278-281.

63. Polesky, A.H., Steitz, T.A., Grindley, N.D.F. and Joyce, C.M. (1990) *J. Biol. Chem*. **265,** 14579-14591.

64. Polesky, A.H., Dahlberg, M.E., Benkovic, S.J., Grindley, N.D. and Joyce C.M. (1992) *J. Biol. Chem*. **267,** 8417-8428**.**

65. Cowart, M., Gibson, K.J., Allen, D.J. and Benkovic, S.J. (1989) *Biochemistry* **28**, 1975-1983.

66. Kuchta, R.D., Cowart, M., Allen, D. and Benkovic, S.J. (1988) *Biochem. Soc. Trans.* **16**, 947-949.

67. Joyce, C.M. (1989) *J. Biol Chem*. **264**, 10858-10866.

68. Lohman, T.M. and Mascotti, D.P. (1992) *Methods Enzymol*. **212**, 400-424.

69. Goeddel, D.V., Yansura, D.G. and Caruthers, M.H. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 3578-3582.

70. Fried, M.G. (1989) *Electrophoresis* **10**, 366-376.

71. Hill, J.J. and Royer, C.A. (1997) *Methods Enzymol*. **278**, 390-416.

72. Heyduk T., Ma, Y., Tang, H. and Ebright, R.H. (1996) *Methods Enzymol*. **274**, 492-503.

73. Record, M.T., Jr, Anderson, C.F. and Lohman, T.M. (1978) *Q. Rev. Biophys*. **11**, 103-178.

74. de Marky, N. and Manning, G.S. (1975) *Biopolymers* **14**, 1407-1422.

75. Manning, G.S. (1978) *Q. Rev. Biophys*, **11**, 179-246.

76. Record, M.T., Jr., Lohman, T.M. and deHaseth, P.L. (1976) *J. Mol. Biol*. **107**, 145-158.

77. O'Brien, R., DeDecker, B., Fleming, K., Sigler, P. B. and Ladbury, J. E., (1998) *J. Mol. Biol.* **279**, 117-125.

78. Bergqvist, S., O'Brien, R. and Ladbury, J.E. (2001) *Biochemistry* **40**, 2419-2425.

79. Wyman, J. (1964) *Adv. Protein Chem*. **19**, 223-226.

80. Record, M.T., Ha, J.H. and Fisher, M.A. (1991) *Methods Enzymol*, **208**, 291-343.

81. Ha, J.H., Capp, M.W., Hohenwalter, M.D., Baskerville, M. and Record, M.T. Jr. (1992) *J. Mol. Biol*. **228**, 252-264.

82. Anderson, C.F., Record, M.T. Jr. and Hart, P.A. (1978) *Biophys. Chem*. **7**, 301-316.

83. Bleam et al (1980) *Proc. Natl. Acad. Sci. U.S.A*. **77**, 3085-3089.

84. Barkley, M.D., Lewis, P.A. and Sullivan, G.E. (1981) *Biochemistry* **20**, 3842-3851

85. deHaseth, P.L., Lohman, T.M. and Record, M.T. (1977) *Biochemistry* **16**, 4783-4790.

86. Overman, L.B., Bujalowski, W. and Lohman, T.M. (1988) *Biochemistry*, **27**, 456-471.

87. Menetski, J.P., Varghese, A. and Kowalczykowski, S.C. (1992) *J. Biol. Chem*. **267**, 10400-10404.

88. Ha, J.H., Spolar, R.S. and Record, M.T. Jr. (1989) *J. Mol.Biol.* **209**, 801-816.

89. Ladbury, J. E., Wright, J. G., Sturtevant, J. M. and Sigler, P. B. (1994) *J. Mol. Biol.* **238**, 669-681.

90. Merabet, E. and Ackers, G. K., (1995) *Biochemistry* **34**, 8554-8563.

91. Jin, L., Yang, J. and Carey, J. (1993) *Biochemistry* **32**, 7302-7309.

92. Lundback, T. Cairns, C., Gustafsson, J.A., Carlstedt-Duke, J. and Hard, T. (1993) *Biochemistry* **32**, 5074-5082.

93. Petri, V., Hsieh, M. and Brenowitz, M. (1995) *Biochemistry* **34**, 9977-9984.

94. Morton, C.J. and Ladbury, J.E. (1996) *Prot. Science* **5**, 2115-2118.

95. Berger, C., Jelesarov, I. and Bosshard, H.R. (1996) *Biochemistry* **35**, 14984-14991.

96. Lipps, G., Stegert, M. and Krauss, G. (2001) *Nuclec Acids Research*, **29**, 904-913.

97. Jen-Jacobson, L., Engler, L.E., Ames, J. T., Kurpiewski, M.R. and Grigorescu, A. (2000) *Supramolecular Chemistry* **12**, 143-160.

98. Sturtevant, J.M. (1977) *Proc. Natl. Acad. Sci. U.S.A*. **74**, 2236-2240.

99. Robertson, A.D. and Murphy, K.P. (1997) *Chem. Rev*. **97**, 1251-1267.

100. Spolar, R.S., Livingstone, J.R. and Record, M.T. (1992) Biochemistry 31, 3947-3955.

101. Murphy, K.P. and Freire, E. (1992) *Adv. Prot. Chem*. **43**, 313-361.

102. Makhatadze, G.I. and Privalov, P.L. (1995) *Adv. Prot. Chem*. **47,** 307-425.

103. Myers, J. K., Pace, C. N. and Scholtz, J. M. (1995) *Protein Sci*. **4**, 2138-2148.

104. Spolar, R. S. and Record Jr., M.T. (1994) *Science*, **263**, 777-784.

105. Takeda, Y., Ross, P.D. and Mudd, C.P. (1992) *Proc Natl Acad Sci U S A* **89**, 8180-8184.

106. Heyduk, E., Baichoo, N. and Heyduk, T. (2001) *J. Biol. Chem*. **276**, 44598-44603.

107. Ferrari, M.E. and Lohman, T.M. (1994) *Biochemistry*. **33**, 12896-12910.

108. Lawyer, F.C., Stoffel, S., Saiki, R.K., Chang, S.Y., Landre, P.A., Abramson, R.D. and Gelfand, D.H. (1993) *PCR Methods Appl.* **2**, 275-287.

109. Hogrefe, H.H., Cline, J., Lovejoy, A.E. and Nielson, K.B. (2001) *Methods Enzymol*. **334**, 91-116.

110. Turner, R.M., Grindley, N.D.F. and Joyce, C.M. (2003) *Biochemistry* **42**, 2373-2385.

111. Delagoutte, E. and von Hippel, P.H. (2003) *J. Biol. Chem*. **278**, 25435-25447.

CHAPTER 2

PREPARATION OF PROTEINS EXAMINED IN THIS STUDY

Introduction

The proteins examined for this thermodynamic study are the full length Taq DNA polymerase, the Klentaq fragment of Taq DNA polymerase (1) and the exonuclease deficient D424A mutant of Klenow fragment (2) of *E.coli* Pol I. Mutants of Klenow polymerase lacking any measurable exonuclease activity (2) have been widely used for structural and functional studies of the polymerase domain in order to minimize DNA degradation; the D424A mutant being the most predominantly used exonuclease deficient mutant of Klenow polymerase (KF exo-). The clone for KFexo- was a gift from Catherine Joyce at Yale University. The clone of Klentaq, constructed by Wayne Barnes at Washington University (1), was obtained from the American Type Culture Collections (ATCC) (Manassas, VA). Both KFexo- and Klentaq were expressed and purified using published procedures (3,4) with slight modifications as described below. There are a number of previously published or patented descriptions of the purification of full length Taq DNA polymerase (4-17). However, in a majority of these procedures, the purpose of production of Taq DNA polymerase is for use in PCR, where high yield and full purification of the protein are not necessary. Our studies are aimed at biophysical characterization of the protein requiring large amounts of highly purified proteins. As such, I have constructed an over-expression clone (pKD-Taq3) of full length Taq DNA polymerase that allowed us to obtain adequate amount of protein for our studies. The sequencing of the construct confirmed the presence of correct Taq DNA polymerase coding sequence in the resulting plasmid. I also designed a more rigorous purification of

the protein using significant modifications of other published procedures for the purification of Taq DNA polymerase. Also, no surfactants have been used in the purification, since their effects on the polymerase remain uncharacterized. The activity of the enzyme in PCR (determined by the yield from PCR) was comparable to the commercial Taq DNA polymerase preparations from Fisher, Promega and Gibco.

Experimental Procedures

Cloning of the Taq Polymerase Gene

Amplification of the Taq DNA Polymerase Gene

The Taq polymerase gene was amplified using the primers K1 and K2 as shown in Figure 2.1. In order to insert restriction sites at the two ends of the polymerase gene, primer K1 contained an extension harboring an EcoR1 site, while K2 contained an extension coding for a BglII site. The primers were designed following the procedure of *Engelke, et al*., *1990* (9). A commercial high fidelity thermostable DNA polymerase (Accuzyme, Midwest Scientific, St. Louis, MO, USA), possessing 5'-3' polymerase and 3'-5' proofreading activities, was used in order to prevent the introduction of any mutations during PCR. The amplification was performed in a 50ul reaction volume consisting of 1X reaction buffer (10mM Tris-HCl, 50mM KCl, pH 8.3), 2mM $MgCl_2$, 0.2mM each of the dNTP's, 1uM of each of the primers, 0.2ng of the template DNA and 5U of commercial high fidelity thermostable DNA polymerase, using an automated PCR thermocycler (GeneAmp, PCR System 9700, PE Applied Biosystems). The reaction condition was comprised of an initial heating up to 94^{o}C for 5min for the separation of the template strands followed by 30 consecutive cycles, each of which consisted of 1min at 94^{o}C, 1min at 55^{o}C for annealing of the primers to the template DNA and 3min at 72^{o}C

for DNA extension by Taq polymerase. In the last cycle, the DNA extension step at 72^{O}C was carried out for an additional 10min and the reaction mixture was cooled down to 4^{o}C. The amplified DNA was purified using a QIAGEN PCR purification kit.

Cloning of the Taq Gene in Over-Expression Plasmids:

A) Cloning into pTrc 99A (Amersham) Expression Plasmid.

The amplified gene was digested with EcoR1 and BglII restriction enzymes (Promega). The pTrc 99A (expression plasmid, Amersham) was digested with EcoR1 and BamHI, and treated with alkaline phosphatase. Each of the above reactions was followed by a purification step (using QIAGEN PCR purification kit) to remove all the proteins. The amplified Taq gene was cloned into the pTrc 99A vector using T4 DNA ligase (Fig. 2.1). The resulting clone (pKD-Taq2) was used to transform the JM109 overexpression strains of *E.coli*. However, the induction of the pKD-Taq2 plasmid produced only small quantities of the protein, requiring subcloning of the gene into another expression plasmid, pET-15b (Novagen).

B) Subcloning of the Taq gene in pET-15b Expression Plasmid.

The Taq gene was isolated from the pKD-Taq2 by restriction digestion with Nco1 and Sal1 and subligated into a pET-15b expression plasmid using the Nco1 and Xho1 restriction sites (Fig.2.1). The His-tag region in the pET-15b vector was purposely removed during the cloning procedure to avoid the presence of any additional amino acid sequence in the Taq DNA polymerase expressed from this clone. The resulting clone (pKD-Taq3) was transformed into the BL21(DE3) expression strain of *E.coli* (Novagen) and adequate amounts of Taq DNA polymerase were obtained from this construct following IPTG induction.

Sequencing of the pKD-Taq3 Clone:

The Taq polymerase gene in the pKD-Taq3 clone was sequenced using the following primers, Each set of forward and reverse primers covered approximately 20-40 bases in the Taq DNA sequence. Both strands were sequenced in each segment.

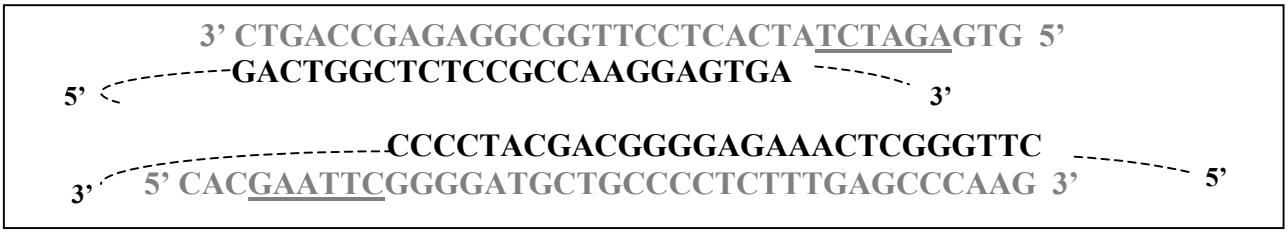
TP1: 5'CACGAATTCGGGGATGCTGC 3' (forward)
TP2: 5'GGGAGCCTGGAAGCCCTCCT 3' (forward)
TP3: 5'TCCTTTGGCTTTACCGGAG 3' (forward)
TP4: 5'GGAGACCGCCAGCTGGATGT 3' (forward)
TP5: 5'GGTGGTATCACTCCTTGGCG 3' (reverse)
TP6: 5'ACATCCAGCTGGCGGTCTCC3' (reverse)
TP7: 5'CTCCCGGTAAAGCCAAAGGA 3' (reverse)
TP8: 5'AGGAGGGCTTCCAGGCTCCC 3' (reverse)
TP9: 5'TGCTCCTCGCCTACCTCCTGGACCCTTCCAACACCACCCC 3'(forward)
TP10: 5'AGGTTGAAGGGGTGGCCGGCCAGGCGGAAGACCTCGGCCT3'(reverse)
TP11: 5'ATGCCCGTCCAGGGCACCGCCGCCGACCTC 3' (forward)
TP12: 5'TCGTCCGCCTCGTAGCCCGGGACCTCGAGG3' (reverse)

The sequencing reaction was kindly performed by the GeneLab facility, School of Veterinary Medicine (Louisiana State University). The sequenced DNA fragments were blasted against the wild type Taq polymerase gene sequence using the program Blastn (ver. 2.2.1) available on the NCBI website (*www.ncbi.nlm.nih.gov*). The sequenced DNA matched the wild type DNA sequence, thereby confirming that the sequence of the Taq polymerase gene in the pKD-Taq3 clone was identical to the wild type DNA polymerase I gene from *Thermus aquaticus.*

Expression and Purification of the Taq DNA Polymerase

The transformed BL21 (DE3) cells were grown on agar plates made with rich media (10g yeast extract, 25g tryptone, 10g M9 salts and 10g glucose per liter of media) at pH 8.5 containing 200ug/ml ampicillin. It important to note that the addition of glucose and a relatively high concentration of ampicillin in the growth media, as well as maintaining the pH at 8.5, all helped toward stability of the plasmid in the host strain. For this reason

Taq Pol I gene amplified with forward and reverse primers incorporating EcoR1 and Bgl II restriction sites, respectively - light gray oligos are the amplification primers, underlined sequences are the EcoR1 and Bgl II cut sites



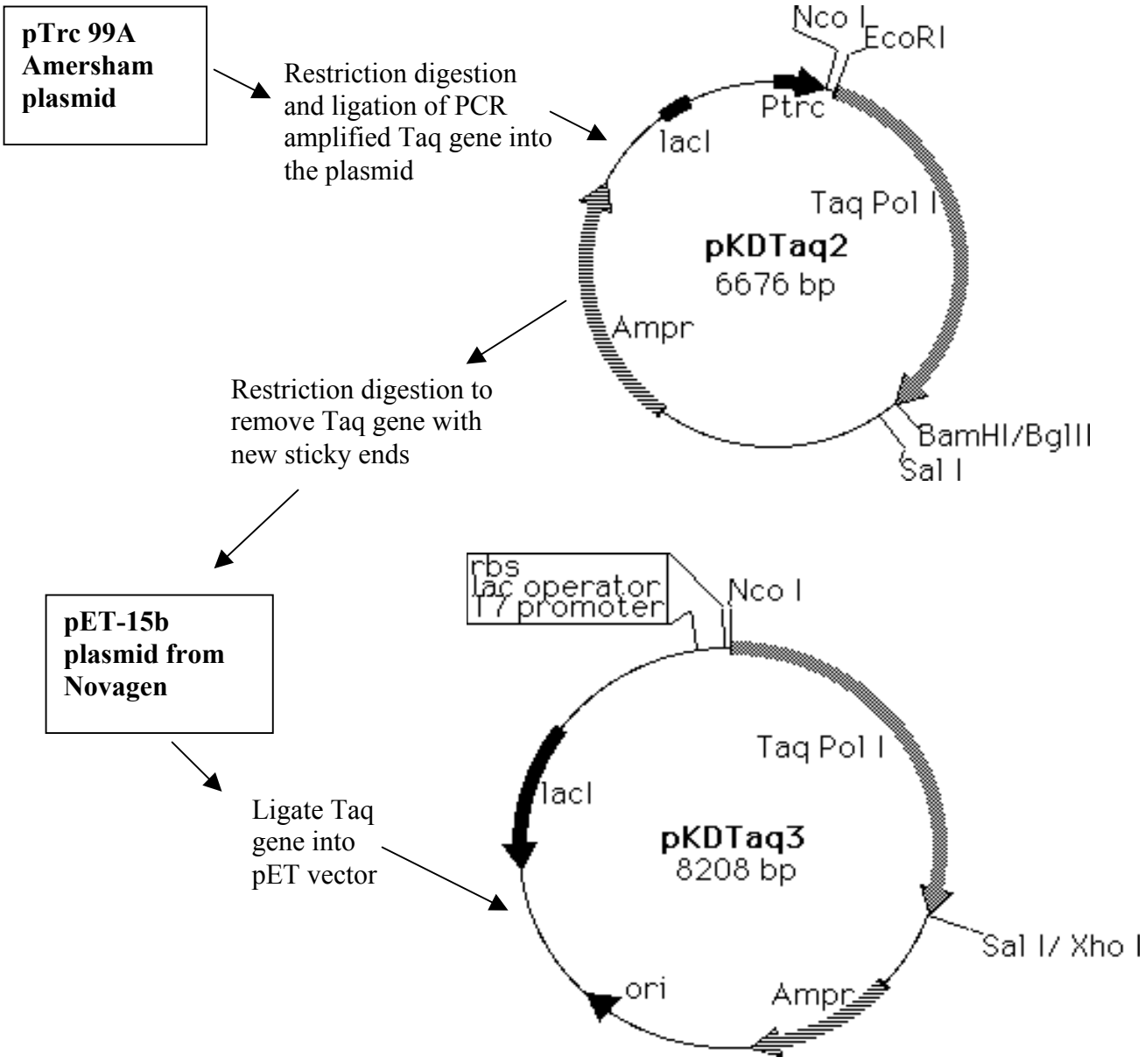


Figure 2.1. Cloning scheme for Taq DNA Polymerase gene.

it is also recommended that a fresh transformation be performed before each purification rather than using frozen cells that had been previously transformed. The transformed cells are grown on the plates for no more than 8-10 hours so that colonies are picked at a very early stage. The cells are then grown in a 5ml liquid media (same as used for the plates, without the agar) followed by inoculation into 1liter media for the expression of the protein. When the cells reached an OD_{550} of 1.2 they were induced with 5mM IPTG and allowed to grow for another 14 hours before harvesting for purification of the protein.

The cell pellets (typically 19-20g/liter culture) were solubilized in TMND buffer (50mM Tris-Cl, 10mM $MgCl_2$, 16mM $(NH_4)_2SO_4$, 50mM Dextrose and 250mM KCl, pH 7.9). Crude cell extract was then prepared by lysis of cell pellets with lysozyme (4mg/ml) for 15min at room temperature followed by heating at 75^{o}C for 1hour. The cell extract was transferred to ice immediately and allowed to cool down for 20mins. The solution was then centrifuged at 6500g in a Beckman centrifuge for 40min and the supernatant was subjected to PEI (polyethylene imine) precipitation to remove the contaminating nucleic acids. For this, 5% neutral PEI solution was gradually added to the supernatant, until there is no further precipitation with the addition of PEI. This step was performed on ice. Generally, a final concentration of 0.3% PEI is sufficient to remove all the nucleic acids from the preparation. The solution was then centrifuged at 6500g and the supernatant was applied to a Bio-Rex 70 cation-exchange column pre-equilibrated with KTA buffer (20mM Tris, 22mM $(NH_4)_2SO_4$, 1mM DTT, 0.1mM EDTA, 10% glycerol, pH 7.9). The flow-through was dialyzed in KTA buffer and applied to a Heparin-agarose column in the same buffer. Taq DNA polymerase was eluted with a 22-270mM

$(NH_4)_2SO_4$ gradient. The eluent was applied to a second Bio-Rex 70 column in KTA at pH 8.8, and the flow-through contained the purified protein (Fig. 2.2).

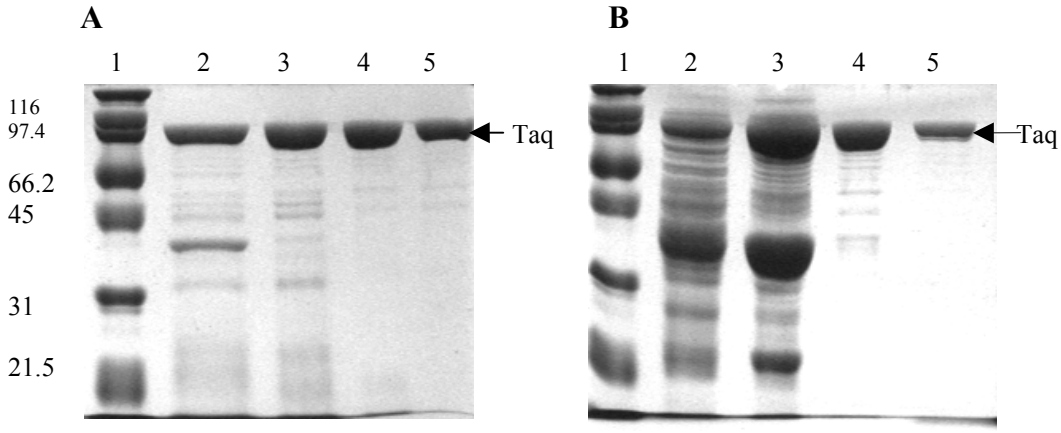


Figure 2.2. A) Purification of Taq DNA Polymerase I with initial heat step. Lanes 1: Molecular weight markers (in kDa), 2: Cell lysate after heating at 75°C for 1hr, 3: After Bio-Rex 70 column at pH 7.9, 4: After Heparin column at pH 7.9, 5: After Bio-Rex 70 column at pH 8.8. B) Purification of Taq DNA Polymerase I without a heat step. Lanes 1: Molecular weight marker (in kDa), 2: After Bio-Rex 70 column, 3: 40% Ammonium sulfate pellet, 4: After Heparin column at pH 7.9, 5: After Hydroxyapatite column at pH 7.9

In addition, we also assayed for the potential damaging effect, if any, of the 75°C heat step included in the purification procedure. For this we purified the protein without the heat step (Fig 2.2). The first Bio-Rex 70 flow-through was subjected to a 40% $(NH_4)_2SO_4$ precipitation. The pellet from the 40% $(NH_4)_2SO_4$ cut was dialyzed in KTA buffer to remove all the excess salt and applied to a Heparin-agarose column in the same buffer. The Taq DNA polymerase eluted from the column was dialyzed in phosphate buffer (10mM K-PO_4, pH 7.5) and loaded onto a hydroxyapatite column pre-equilibrated with the phosphate buffer. Taq DNA polymerase was eluted with a 0-500mM KCl gradient (Fig. 2.2). However, this "cold-prep" polymerase behaved identically in all our

assays as the normal "heat-prepared" polymerase, and so the heat incubation step was retained in the final purification protocol.

Expression and Purification of Klentaq and KFexo- DNA Polymerases

Klentaq polymerase was purified as published (1) with the following modifications: no surfactant was used during purification, the ammonium sulfate precipitation step was omitted, and a second Bio-Rex 70 column, at pH 9.1, was added subsequent to the heparin column. The flow through of the second Bio-Rex 70 column contains the purified protein.

KFexo- (plasmid pXS106) was transformed into *E.coli* expression strain CJ376 and grown and induced as described previously (18). The protein was purified as described was full-length Pol I (3) with omission of the DEAE-cellulose column. No surfactant was used during purification or storage of any of the polymerases. Protein concentrations were measured using the Bradford method (19).

References

1. Barnes, W.M. (1992) *Gene* **112**, 29-35.

2. Derbyshire, V., Freemont, P.S., Sanderson, M.R., Beese, L., Friedman, J.M., Joyce, C.M. and Steitz, T.A. (1988) *Science* **240**, 199-201.

3. Minkley, E.G., Jr., Leney, A.T., Bodner, J.B., Panicker, M.M. and Brown, W.E. (1984) *J. Biol. Chem*. **258**, 10386-10392.

4. Barnes, W.M. (1995) (July 25, 1995) US Patent 5,436,149.

5. Chien, A., Edgar, D.B. and Trela, J.M. (1976) *J. Bacteriol.* **127**:1550-1557.

6. Kaledin, A.S., Sliusarenko, A.G. and Gorodetskii, S.I. (1980) *Biokhimiia* **45**, 644-651.

7. Lawyer, F.C., Stoffel, S., Saiki, R.K., Myambo, K., Drummond, R. and Gelfand, D.H. (1989) *J. Biol. Chem*. **264**, 6427-6437.

8. Lawyer, F.C., Stoffel, S., Saiki, R.K., Chang, S.Y., Landre, P.A., Abramson, R.D. and Gelfand, D.H. (1993) *PCR Methods Appl.* **2**, 275-287.

9. Engelke D.R., Krikos, A., Bruck, M.E. and Ginsburg, D. (1990) *Anal. Biochem*. **191**, 396-400.

10. Gelfand, D.H., Stoffel, S., Lawyer, F.C. and Saiki, R.K. (1989) (December 26, 1989) US Patent number 4,889,818.

11. Gelfand, D.H., Stoffel, S., Lawyer, F.C. and Saiki, R.K. (1992) (January 7, 1992) US Patent number 5,079,352.

12. Grimm, E. and Arbuthnot, P. (1995) *Nucleic Acids Res*. **23**, 4518-4519.

13. Pluthero, F.G. (1993) *Nucleic Acids Res*. **21**, 4850-4851.

14. Leelayuwat, C., Srisuk, T., Paechaiyaphum, R., Limpaiboon, T., Romphruk, A. and Romphruk, A. (1997) *J. Med. Assoc. Thail*. **80,** Suppl. 1, 129-137.

15. Dabrowski, S. and Kur, J. (1998) *Acta. Biochim. Pol*. **45**, 661-667.

16. Nord, K., Gunneriusson, E, Uhlen, M. and Nygren, P.A. (2000) *J. Biotechnol*. **80,** 45-54.

17. Sullivan, M.A. (2000) (July 4, 2000) U.S. Patent 6,083,686.

18. Joyce, C.M. and Derbyshire, V. (1995) *Methods Enzymol*. **262**, 3-13.

19. Bradford, M.M. (1976) *Anal. Biochem*. **72**, 248-254.

CHAPTER 3

SALT DEPENDENCE OF DNA BINDING BY *THERMUS AQUATICUS* AND *ESCHERICHIA COLI* DNA POLYMERASES*

Introduction

The non-covalent driving forces that lead to stable protein-DNA complexes are strongly influenced by the solution environment (salt concentration and type, temperature, pH, etc.). As a result of this dependence on the solution conditions it is impossible to understand the forces that drive these interactions based solely on structural considerations. Rather, the functional properties (thermodynamics and kinetics) of these interactions must be investigated as a function of solution conditions in order to understand the origins of the stability of the complexes. Models and mechanistic explanations of function in the Pol I polymerase family are often extrapolated from one family member to another. These extrapolations are well justified given the structural similarities among the family members as revealed by the extensive series of recent co-crystal structures of different family members (for review, see Refs. 1 and 2) as well as sequence similarities among the active sites (1). Assays of function and mutagenesis studies, however, more often reveal subtle and not so subtle differences among the family members (e.g. Refs. 3 and 4) but directly comparative functional studies are relatively scarce. In this study, we have characterized the basic binding equilibria of full length Taq, Klentaq, and Klenow polymerases to DNA using a fluorescence anisotropy assay.

---

* Reprinted with permission from *The Journal of Biological Chemistry,* **Vol. 278(8),** 5694-5701.

We have further characterized the KCl and $MgCl_2$ dependencies of DNA binding by the different polymerases, and have quantitated the linked ion releases upon DNA binding for all three polymerases. The results show that these two species of polymerase, when binding to the same DNA, differ in their salt "sensitivities", their intrinsic affinities, their linked ion releases, and their need for bound divalent cations.

Experimental Procedures

Proteins

The proteins examined are full length Taq DNA polymerase, the Klentaq fragment of Taq polymerase (5), and the D424A mutant of Klenow polymerase (6). The proteins were purified as described in Chapter 2 of this dissertation. No surfactant was used during purification or storage of any of the polymerases. Protein concentration was measured using the Bradford method (7).

DNA

Stoichiometric and equilibrium DNA binding experiments were performed with the following primer-template sets:

13/20mer:

```
5′TCGCAGCCGTCCA 3′
3′AGCGTCGGCAGGTTCCCAAA 5′
```

63/70mer:

```
5′ TACGCAGCGTACATGCTCGTGACTGGGATAACCGTGCCGTTTGCCGACTTTCGCAGCCGTCCA 3′
3′ ATGCGTCGCATGTACGAGCACTGACCCTATTGGCACGGCAAACGGCTGAAAGCGTCGGCAGGTTCCCAAA 5′
```

The 13/20mer primer template set used is the same as that used for kinetic studies of Klenow DNA binding by Benkovic and associates (8). The longer primer-template pair (63/70mer) was designed for use at higher temperatures. The 63/70mer uses the same putative binding region sequence as the 13/20mer, with added "random" sequence

selected from the same region of the M13 bacteriophage sequence. Labeled and unlabeled DNA oligonucleotides were purchased from Integrated DNA Technologies Inc. Fluorescently labeled DNA was labeled at the 5'end of the primer strand with Rhodamine-X (ROX) and was purchased directly from Integrated DNA Technologies.

Fluorescence Anisotropy Assay

ROX-labeled DNA was titrated with increasing concentrations of protein and binding was monitored using the anisotropy signal change as protein-DNA complex is formed (9,10). Fluorescence anisotropy measurements were performed using a FluoroMax-2 fluorometer equipped with an automated polarizer and regulated at the indicated temperatures. The excitation and emission wavelengths were 583nm and 605nm respectively, with 8nm band-pass and integration time of 10 seconds. For all the equilibrium titrations the DNA concentration used was 1nM. In all the experiments the protein was titrated into fluorescently labeled DNA, with the total [DNA] << K_d. After each addition the sample was equilibrated at the required temperature for 8 min and anisotropy was measured. All titrations were performed in 10mM Tris, pH 7.9 buffers at the indicated salt (KCl and $MgCl_2$) concentrations and temperature. The pH was adjusted by mixing Tris base and Tris-HCl. KCl and $MgCl_2$ concentrations were varied across the widest possible range for each polymerase. Limiting titrations at low KCl concentrations were those where the K_d approached the total [DNA]. Limiting values at high KCl concentrations were the point where well-behaved, reproducible isotherms could no longer be obtained. For assays examining the effects of EDTA on binding, the titrations were carried out in 10mM Tris, 10mM EDTA, pH 7.9, with KCl concentrations of 50mM for Klentaq and 300mM for Klenow. Prior to measurement of the binding in the presence of EDTA the proteins were extensively dialyzed against this same buffer.

Data Analysis

Stoichiometric binding curves were fit to the equation,

$$\Delta A=\Delta A_T/2D_T\{(E_T + D_T + K_d)-[(E_T + D_T + K_d)^2 - 4E_TD_T]^{1/2}\} \quad \textbf{(Eq. 3.1)}$$

where, ΔA is the change in fluorescence anisotropy, ΔA_T is the total change in anisotropy, E_T is the total polymerase concentration at each point in the titration, D_T is the total DNA concentration and K_d is the dissociation constant for polymerase-DNA binding (9,11). When used to analyze stoichiometric binding curves, both D_T and K_d are allowed to vary during the non-linear regression, and the binding stoichiometry was determined as the ratio of the fitted D_T and the known D_T.

Equilibrium binding curves were fit both to the equation above and to a simplified standard single-site isotherm usable when the [DNA] « K_d, and which thus assumes effective equality between the free polymerase concentration and the total polymerase concentration,

$$\Delta A = \{\Delta A_T(E_T /K_d\}/\{1+ E_T/K_d\} \quad \textbf{(Eq.3.2)}$$

where, ΔA is the change in fluorescence anisotropy, ΔA_T is the total change in anisotropy, E_T is the total polymerase concentration at each point in the titration, and K_d is the dissociation constant for polymerase-DNA binding. Both equations have strengths and weaknesses for fitting equilibrium binding data. The fitted K_d in Equation 3.1 is quite sensitive to the exact fixed total DNA concentration and any uncertainty in the determination of this concentration, whereas in Equation 3.2 the actual difference between the total and the free polymerase concentrations propagates directly into the fitted K_d. Fits to the titrations with the very tightest (lowest) K_d values in each salt series in this study gave ≤10% variation in the fitted K_d values using Equations 3.1 and 3.2. All other titrations in all salt series gave, within error, identical fitted K_d values with

both equations. Values and errors reported in the tables are those returned from fits to Equation 3.1. All non-linear fitting was performed using the program KaleidaGraph (Synergy Software). Linked ion release upon binding of the polymerases to DNA was calculated using a basic linkage relationship (12-15), e.g. for KCl the relationship is the following.

$$\{\partial \ln 1/K_d\}/\{\partial \ln[KCl]\} = \Delta n_{ions} = \Delta nK^+ + \Delta nCl^- \quad \textbf{(Eq. 3.3)}$$

Thus the slope of a plot of ln $1/K_d$ *versus* ln [KCl] will be equivalent to the net number of ions (Δn_{ions}) that are bound or released when the protein-DNA complex is formed, where Δn_{ions} is the net sum of the binding and release of both the anions and cations ($\Delta nK^+ + \Delta nCl^-$) (13-15).

Results

In order to obtain adequate amounts of Taq polymerase to perform biophysical studies, we subcloned the gene for Taq polymerase into a overexpressing pET vector (Novagen) using modification of the strategy and amplification primers originally described by Engelke, et. al. (16) (See Chapter 2). For routine use in PCR, full purification of the polymerase is not required, and the majority of the published or patented descriptions of Taq purification are aimed at producing Taq for use in PCR, whereas our procedures have been designed for producing Taq for use in biochemical and biophysical studies. As such, no additional tagged sequences (such as his-tags) have been added to the polymerase. Also, no surfactants have been used in the purification, since their effects on the polymerase remain uncharacterized. In addition, we assayed for potential damaging effects of the ubiquitous 70^{o}C heat incubation step included in published and patented Taq purification procedures. Full purification of the polymerase without the heat step was performed, however this "cold prep" polymerase behaved

identically in all of our assays as the normal "heat prepared" polymerase (data not shown), and so the heat incubation step was retained in the final purification protocol.

Binding Stoichiometry of the Polymerases:

Stoichiometric titrations of each of the polymerases versus DNA were performed using fluorescence anisotropy and results are shown in Figure 3.1 for Taq polymerase. Polymerases were titrated into Rhodamine-X-(ROX)-labeled DNA at high concentrations of DNA ([DNA] >> K_d). The anisotropy of the DNA increases as protein binds due to the decreasing rate of molecular rotation of the DNA in the complex. The stoichiometry of binding was determined by fitting the data to Equation 3.1 as described under "Experimental Procedures". The ratio of bound protein to DNA at saturation is 0.97 for Taq, 1.04 for Klentaq and 0.82-0.95 for Klenow binding to the 13/20mer primer-template pair. The binding stoichiometry to the longer 63/70mer primer-template pair was 1.1 for Taq and 1.15 for Klentaq. As a control, it was determined if unlabeled DNA could effectively compete with fluorescently labeled DNA for binding to the polymerase. The second curve in the stoichiometric titration plots in Figure 3.1 shows a stoichiometric titration of protein versus double the amount of DNA, where half the DNA is ROX-labeled and half is not. Only the fluorescently labeled DNA is anisotropically visible, so the displacement and the apparent doubling of the stoichiometric value for this second curve indicates approximately equally effective binding of both the labeled and unlabeled DNA to the polymerase. Unlabeled DNA effectively competed with ROX-labeled DNA binding to all three polymerases. It should be noted that DNA labeled with fluorescein, which was examined first, is not effectively competed by unlabeled DNA for binding (data not shown), underscoring the importance of the unlabeled competitive control.

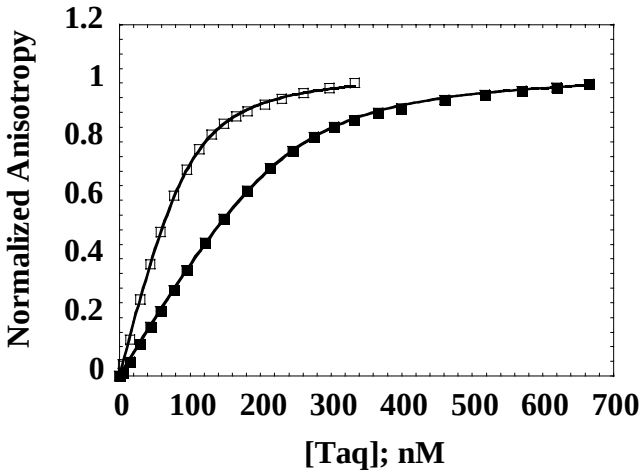


Figure 3.1. Determination of the binding stoichiometry for Taq polymerase binding to DNA. Shown are fluorescence anisotropy-monitored stoichiometric titrations of Taq polymerase into 100nM ROX labeled DNA (□), and an unlabeled DNA competition control titration containing 100nM ROX labeled plus 100nM unlabeled DNA (■). The titrations were performed in 10mM Tris, 5mM $MgCl_2$, 50mM KCl pH7.9 at 25^{O}C with 13/20mer DNA. The lines shown are the fits to Eq. 3.1.Equivalent stoichiometric titrations were carried out for Klentaq and Klenow polymerases, and for Taq polymerase binding to 63/70mer DNA at 25^{o}C and 60^{o}C (data not shown).

KCl Dependence of DNA Binding.

Figure 3.2 shows representative fluorescence anisotropic titrations of the binding of Klenow and Klentaq polymerases to identical pieces of DNA at several different KCl concentrations at 25^{o}C. Each titration curve fits well to a single-site binding isotherm, and it can be seen from the precision of the data that even modest shifts in K_d can be readily quantitated. Figure 3.3 shows the thermodynamic linkage plots for DNA binding as a function of KCl concentration ($\partial \ln 1/K_d$ versus $\partial \ln[salt]$) for the polymerases in the presence and absence of 5mM $MgCl_2$. The negative slopes of the linkage plots are indicative of net ion release upon formation of the protein-DNA complex (12-15). As observed for most DNA binding proteins, DNA binding is linked to ion release for all the

polymerases. Full length Taq polymerase and the Klentaq fragment behave essentially identically. The exact linked ion releases are reported in Table 3.1.

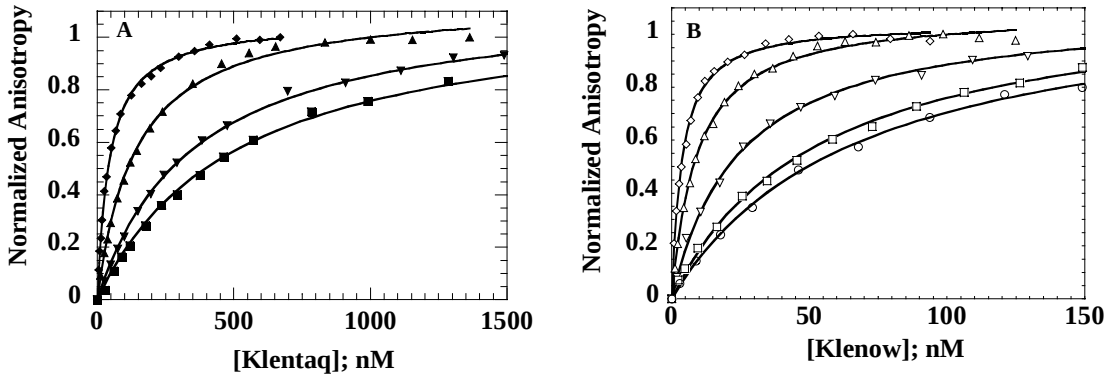


Figure 3.2. Representative equilibrium binding isotherms for the binding of Klentaq and Klenow polymerases to 13/20mer DNA (1nM concentration in all titrations) at various [KCl] concentrations at 25°C. Binding was monitored using fluorescence anisotropy. A, equilibrium titrations of Klentaq at KCl concentrations of 75mM (♦), 125mM (▲), 150mM (▼), and 175mM (■). B, equilibrium titrations of Klenow at KCl concentrations of 250mM (◇), 300mM (Δ), 400mM (∇), 450mM (□), and 500mM (○).

Several points are notable about these data. The most striking observation is the significant difference in salt ranges where sub-micromolar binding occurs. DNA binding across the ~10-300 nM K_d range occurs at KCl concentrations about an order of magnitude higher for Klenow than for Taq/Klentaq. This means that under similar solution conditions the binding of Klenow to DNA is much tighter than that of Taq/Klentaq. Both species of polymerase were titrated over the widest possible KCl concentration ranges that produced acceptable and analyzable titration curves. Although the two [KCl] ranges do not overlap, by extrapolation the binding of Klenow is, on average, about 3 kcal/mol (~150X) tighter than the binding of Taq/Klentaq to the same DNA at similar salt concentrations. This difference is also reflected in the extrapolation of the linkage plots to 1M salt, which provides an estimate of the non-electrostatic components of the binding interaction (13-15). At 1M salt the binding free energy for

Klenow to DNA is also ~3 kcal/mole tighter than the $\Delta G_{binding}$ of Taq/Klentaq to the same DNA (Table 3.1).

In addition to the intrinsic affinity difference between the *E. coli* and Taq polymerases, the slopes of the linkage plots with KCl indicate that the binding of Klenow to DNA results in the release of 1.4 more ions than the binding of Taq/Klentaq to the same DNA, both in the presence and absence of $MgCl_2$. This is an approximately 50% increase in the linked ion release for Klenow relative to Taq/Klentaq. For both polymerases, removal of $MgCl_2$ from the buffer increases the reported linked ion release by about 0.6 ions, indicating that part of the ion release in the presence of $MgCl_2$ is due to magnesium ion release.

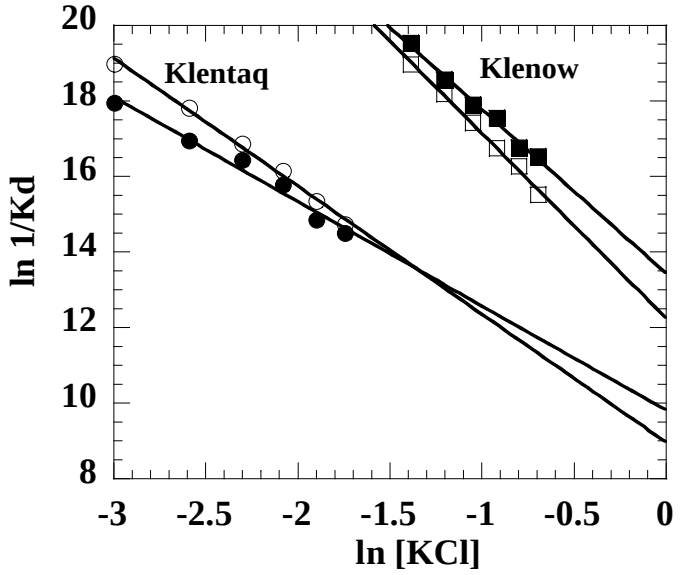


Figure 3.3. KCl linkages ($\partial \ln 1/K_d$ versus $\partial \ln[KCl]$) for the binding of Klentaq and Klenow polymerases to 13/20mer DNA at 25°C. Shown are Klentaq in the presence (●) and absence (○) of 5mM $MgCl_2$, and Klenow in the presence (■) and absence (□) of 5mM $MgCl_2$. The slopes of the plots give the thermodynamic net average number of ions released upon complex formation.

Table 3.1. KCl Dependence of DNA Binding by Klentaq and Klenow with and without $MgCl_2$

All titrations were performed at 25°C in 10mM Tris, pH 7.9, with the listed concentrations of KCl. Titrations performed "with $MgCl_2$" contained 5mM $MgCl_2$.

Klentaq

[KCl] mM	K_d (nM) with $MgCl_2$	Ions released[a]	ΔG (kcal/mol) at 1M KCl	K_d (nM) without $MgCl_2$	Ions released[a]	ΔG (kcal/mol) at 1M KCl
50	16.1 ± 0.3			5.7 ± 0.3		
75	43.5 ± 0.9			18.1 ± 0.7		
100	72.4 ± 2	2.8± 0.2	-5.9	47.2 ± 2	3.4± 0.1	-5.3
125	140.3 ± 5			96.4 ± 4		
150	355 ± 8			215 ± 9		
175	506 ± 15			401 ± 11		

Klenow

[KCl] mM	K_d (nM) with $MgCl_2$	Ions released[a]	ΔG (kcal/mol) at 1M KCl	K_d (nM) without $MgCl_2$	Ions released[a]	ΔG (kcal/mol) at 1M KCl
250	3.3 ± 0.1			5.8 ± 0.2		
300	8.7 ± 0.3			12.8 ± 0.5		
350	16.7 ± 0.7	4.3± 0.2	-8.0	27.3 ± 1.4	4.9± 0.2	-7.4
400	24± 0.8			53.3 ± 2.5		
450	52.3 ± 2			86 ± 5		
500	65.8 ± 2			183.6 ± 12		

[a]Values determined from the slope of the $\partial \ln 1/K_d$ versus $\partial \ln[KCl]$ linkage

$MgCl_2$ Dependence of DNA Binding.

Binding titrations were also performed as a function of $MgCl_2$ at fixed KCl concentrations for Klentaq and Klenow polymerases. Figure 3.4 shows the individual titration curves at different $MgCl_2$ concentrations for the two polymerases, as well as the thermodynamic linkage plots for binding versus [$MgCl_2$]. It is clear that binding of Klentaq is linked to the release of Mg^{+2} across the entire [$MgCl_2$] range, but the [$MgCl_2$] dependence of Klenow DNA binding is not monotonic. The $MgCl_2$ linkage plot for Klentaq (Figure 3.4.C) indicates a linked Mg^{+2} release of 0.8 ions. For Klenow, on the other hand, the $MgCl_2$ linkage plot, and inspection of the titration curves themselves (Figure 3.4.B), both show the absence of a linked Mg^{+2} release (rather, a small uptake) upon binding of Klenow to DNA up until 10mM $MgCl_2$, while above 10mM $MgCl_2$ there is a net release of 1.1 Mg^{+2} upon binding (Table 3.2).

To further investigate the Mg^{+2} requirements for DNA binding by the two polymerases, we EDTA treated Klentaq and Klenow, to remove excess and weakly bound Mg^{+2}, and then assayed DNA binding affinity in the presence of EDTA. In both cases, EDTA significantly decreases the affinity of the polymerases for DNA, as shown in Figure 3.5. This effect is slightly more dramatic for Klenow than for Klentaq (13X reduction of affinity for Klenow versus a 5.5 fold reduction of affinity for Klentaq). Surprisingly, however, removal of EDTA, almost completely restores the original binding affinity of the polymerases (data not shown). These data indicate that free Mg^{+2} is not required for DNA binding by either polymerase, but suggest that some protein bound divalent cations are required for highest affinity binding by both species of polymerase (see Discussion).

Table3.2. $MgCl_2$ Dependence of DNA Binding for Klentaq and Klenow Polymerases

Titrations were performed at 25°C in either 10mM Tris, 50mM KCl, pH 7.9 (Klentaq) or 10mM Tris, 300mM KCl, pH 7.9 (Klenow).

[$MgCl_2$] mM	Klentaq K_d (nM)	Ions released[a]	[$MgCl_2$] mM	Klenow K_d (nM)	Ions released[a]
0	5.7 ± 0.3		0	12.8 ± 0.5	
2	9.2 ± 0.3		2	10.5 ± 0.5	
3	11.2 ± 0.3	0.9± 0.1	5	8.5 ± 0.3	
5	17.6 ± 0.5		10	8.1 ± 0.3[b]	
7	23.5 ± 0.5		15	11.6 ± 0.2[b]	1.2 ± 0.14
10	38.7 ± 2		20	11.7 ± 0.5[b]	
			30	21.3 ± 0.7[b]	
			40	37.4 ± 0.8[b]	
			50	52.4 ± 2[b]	

[a] values determined from the slope of the $\partial \ln 1/K_d$ versus $\partial \ln[KCl]$ linkage.
[b] values used to calculate linked ion release for Klenow

KCl Dependence of Taq Binding at High Temperature:

Thermus aquaticus lives at a temperature optimum of 70-75°C (17). Because it is possible that increased temperature may alter the way Taq polymerase binds DNA, we also assayed for salt dependence of binding at higher temperature. Binding at 60°C required design and use of a longer DNA primer-template pair (63/70mer), since the shorter 13/20mer primer-template had a T_m of ~55°C. Figure 3.6 shows that the salt dependence of binding of Taq polymerase to DNA is very similar, but slightly decreased at 60°C relative to 25°C. K_d values are reported in Table 3.3. Figure 3.6 also shows the KCl dependence for the longer 63/70mer DNA at 25°C, to account for any possible changes due to the differences between the two lengths of DNA.

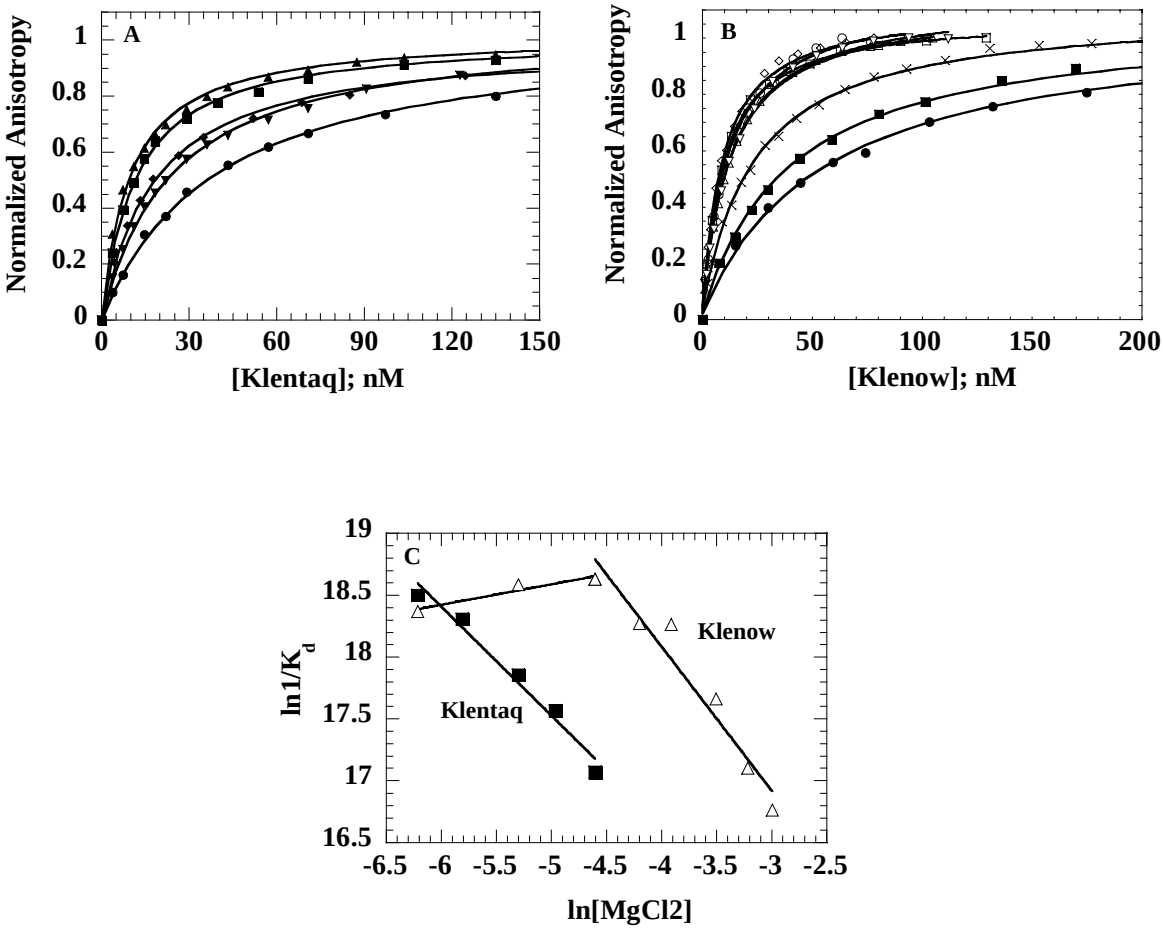


Figure 3.4. $MgCl_2$ dependence of equilibrium binding of Klentaq and Klenow DNA polymerases to 13/20mer DNA (1nM concentration in all titrations) monitored using the fluorescence anisotropy assay at 25°C. A, equilibrium titrations of Klentaq in 10mM Tris, 50mM KCl, pH 7.9, at $MgCl_2$ concentrations of 2mM (▲), 3mM (■), 5mM (♦), 7mM (▼), and 10mM (●). B, equilibrium titrations of Klenow in 10mM Tris, 300mM KCl, pH 7.9, at $MgCl_2$ concentrations of 2mM (○), 5mM (□), 10mM (◇), 15mM (Δ), 20mM (∇), 30mM (x), 40mM (■), and 50mM (●). C, $MgCl_2$ linkages ($\partial \ln 1/K_d$ versus $\partial \ln[MgCl_2]$) for DNA binding of Klentaq (■) and Klenow (Δ).

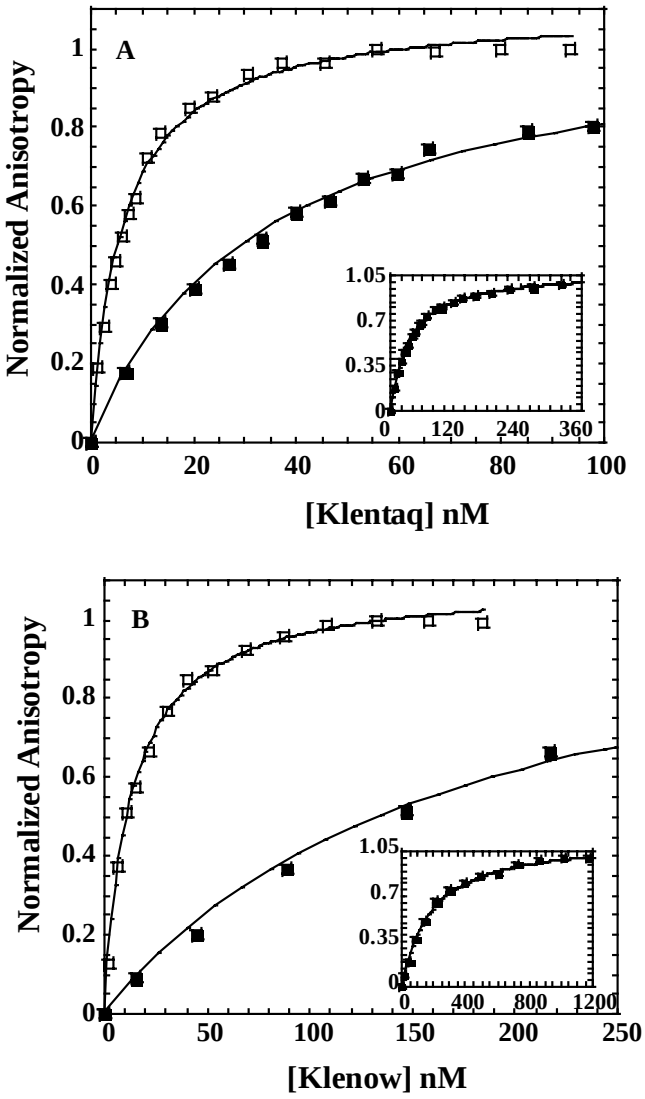


Figure 3.5. Effects of EDTA on the binding of Klentaq (A) and Klenow (B) polymerases to 13/20mer DNA. Equilibrium titrations are shown for both polymerases in the absence of any added $MgCl_2$ (□) and in the presence of 10mM EDTA (■). Binding in the presence of EDTA was always preceded by extensive dialysis of the protein against buffer containing 10mM EDTA. The insets in the figures show the full titrations for the EDTA treated proteins (■).

At 25°C, the binding affinity of Taq for the 63/70mer is somewhat tighter (~3-4X or ~1kcal/mole) than the binding affinity for the 13/20mer, but the $\partial \ln 1/K_d$ versus $\partial \ln[KCl]$ dependencies are the same. It is also interesting to note that Taq binds the 63/70mer DNA with somewhat higher affinity at 25°C than at 60°C. Thermophilic proteins are generally almost catalytically inactive at room temperature, and this is also the case for Taq, which has been shown to have little or no polymerization activity below about 40°C (18). Taq and Klentaq bind DNA quite happily at 25°C, however, and even lower temperatures (See Chapter 4 of this dissertation).

Discussion

KCl Dependence of DNA Binding by the Polymerases.

The linked ion releases for binding of the two different species of polymerases to DNA are both relatively small compared to the range of ion releases generally reported for DNA binding proteins (12-15, 19, 20). Some similar small ion releases that have been reported include the release of 3 ions upon binding of Sso7d to DNA (21) and a 1-2 ion release upon binding of the α-subunit of *E. coli* RNA polymerase to DNA (22). Although the ion release is small for both species of polymerase, the binding of Klenow polymerase results in a larger release of ions than the binding of Klentaq or Taq polymerase. On the scale of the overall ion release for the two polymerases, this difference between Taq/Klentaq and Klenow is somewhat significant, and is indicative of differences in the way the two polymerases interface with DNA. Even though both species of polymerase bind to the same DNA with 1:1 stoichiometry, their different linked ion releases suggest that they bind with different footprints. Below, we discuss this in light of the available crystal structures of the binary complexes of the polymerases bound to primed-template DNA (23,24).

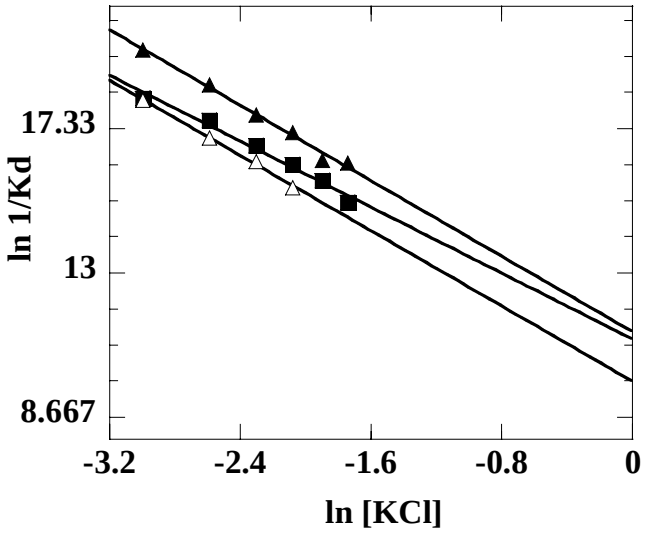


Figure 3.6. KCl linkages for Taq polymerase at different temperatures. $\partial \ln 1/K_d$ versus $\partial \ln[KCl]$ are shown for binding of full length Taq to 1) 13/20mer DNA at 25°C (Δ), 2) 63/70mer DNA at 25°C (▲); and 3) 63/70mer DNA at 60°C (■). All titrations were performed in a base buffer of 10mM Tris, 5mM $MgCl_2$, pH 7.9, with varying KCl concentrations.

Table 3.3. KCl Dependence of DNA Binding for Full Length Taq pPolymerase at Different Temperatures and with Different Lengths of DNA

All titrations were performed in 10mM Tris, 5mM $MgCl_2$, pH 7.9 at the listed concentrations of KCl

Binding to 13/20mer at 25°C

[KCl][a] mM	K_d nM	Ions released[a]	ΔG (kcal/mol) at 1M KCl
50	12.7 ± 0.5		
75	39.7 ± 2.4	2.8 ± 0.12	-5.8
100	80.1 ± 3.3		
125	177 ± 2.6		

(table cont'd)

Binding to 63/70mer at 25°C

[KCl][a] mM	K_d nM	Ions released[a]	ΔG (kcal/mol) at 1M KCl
50	2.8 ± 0.2		
75	8.1 ± 0.4	2.8 ± 0.14	-6.4
100	20 ± 0.5		
125	34 ± 1.6		
150	78 ± 4		
175	83 ± 4		

Binding to 63/70mer at 60°C

[KCl][a] mM	K_d[c] nM	Ions rel[a]	ΔG (kcal/mol) at 1M KCl
50	12.1 ± 0.4		
75	23.2 ± 1	2.5 ± 0.17	-7.3
100	50.0 ± 3.4		
125	89.1 ± 6		
150	142.3 ± 7		
175	277.5 ± 19		

[a] Values determined from the slope of the $\partial \ln 1/K_d$ versus $\partial \ln[KCl]$linkage.

The ion release for Taq polymerase at 60°C is slightly lower than the ion release at 25°C, further distancing the salt dependent behaviors of the two species of polymerase when they are compared at temperatures near their respective optimal growth temperatures. It may be possible that the temperature dependence of the ion release for Taq is more complex than is revealed by looking at only these two temperatures, and that

we have fortuitously chosen two temperatures for Taq where the ion releases are similar. There may exist specific separate solution conditions under which Klentaq and Klenow polymerases might release the same number of ions upon DNA binding, but under the identical solution conditions examined herein, they certainly do not.

Table 3.1 also shows that the $\partial \ln 1/K_d$ versus $\partial \ln[KCl]$ linkage plots for Taq and Klentaq have nearly identical slopes. The Taq and Klentaq $\partial \ln 1/K_d$ versus $\partial \ln[KCl]$ dependencies, shown in Figures 3.4 and 3.6, are nearly co-linear if plotted on the same plot. This finding suggests that in both cases only the polymerase domain of the protein is binding DNA, i.e. that the 5'nuclease domain of full length Taq is not binding DNA under these conditions. If both the 5'nuclease domain and the polymerase domain of the full length enzyme were binding the DNA, one would expect the full length protein to interact with a larger amount of the DNA and hence to exhibit a different ion release.

Intrinsic DNA Binding Affinity and Salt Tolerance Differences Between the Polymerases

One of the most striking differences between the DNA binding properties of Klenow versus Taq/Klentaq polymerases is the shift in salt ranges for sub-micromolar binding. A consequence of this difference is that at any particular salt concentration, the relative affinity of Klenow polymerase for DNA is significantly tighter than that of Taq/Klentaq. This is further suggested by the estimated difference in non-electrostatic binding free energies for the polymerases, estimated from the extrapolation of the KCl dependence data to 1 M salt. Both of these lines of evidence indicate that Klenow binds DNA about 3 kcal/mole tighter than Taq/Klentaq DNA. The data at high temperature for Taq polymerase clearly show that this affinity difference is not due to the fact that we are assaying the binding of Taq/Klentaq to DNA at 25^{o}C. This large difference in relative

binding affinities for the two different species of polymerase, like the difference in linked ion release, further suggests differences in their initial DNA binding modes.

It should be noted that the expression clones for Taq versus Klentaq polymerases in this study are from completely different sources, with no overlap of materials or design. The clone for full length Taq was produced in our laboratory (see Chapter 2), while that for Klentaq was produced by Wayne Barnes (5,25). The fact that the clones from the two independent sources produce Taq and Klentaq proteins that behave almost identically indicates that the decreased DNA binding affinity of Taq/Klentaq is an intrinsic property of the polymerase, and is not the result of some sort of unintentionally introduced mutation. Further, we examined the possibility that the heating step, included in all published Taq/Klentaq purifications, might somehow have partially damaged the protein (see Chapter 2). However, Taq polymerase purified without the heating step behaves identically to the protein purified with the heating step.

It is currently not clear why *E. coli*'s Klenow polymerase binds DNA with so much higher affinity than the equivalent polymerase from *Thermus aquaticus*. Another way to look at the data of Figure 3.3 is to note that the two species of polymerase bind DNA with similar affinities in very different salt concentration ranges. However, neither *E. coli* nor *T. aquaticus* is halophilic. In fact neither will grow well in media containing > 1% NaCl (26, 27). How sensitive their growth and viability are in the 0-1% salt range, however, has not been characterized, so it is possible that there are subtle differences in the salt sensitivities of the two bacteria that could correlate with the salt sensitivity of their Pol I type polymerases. Several other potential explanations for the observed differences may be postulated: 1) There may be other factors in the intracellular environments of the two bacteria that either lower the affinity of the *E. coli* enzyme, or

increase the affinity of the Taq enzyme in vivo, or both (i.e. allosteric regulators, specific anions, other proteins, etc.). Specific anions have been shown to regulate DNA binding affinity in several systems (14, 15, 28). 2) High temperature stability of Taq/Klentaq may be achieved at the cost of the loss of high affinity DNA binding. 3) The two polymerases might show significant differences in sequence specificity, e.g. Taq might bind GC-rich sequences (more common in thermostable bacteria) tighter than mixed DNA sequences. 4) The operational or tolerable functional affinity range for Pol I type polymerases may be quite wide in vivo, such that any affinity in the nanomolar to picomolar range is adequate for appropriate function or easily compensated for in vivo by increased/decreased expression of the particular polymerase. 5) These Pol I type polymerases from *E. coli* and *T. aquaticus*, although widely considered homologues, may not be completely functionally equivalent in the two bacteria, and therefore might not be expected to have similar DNA affinities. 6) DNA may simply bind to the two polymerases in different ways. This point requires expansion. It has long been known that Taq DNA polymerase is devoid of 3'exonuclease (or proofreading) activity. The proofreading activity of Klenow polymerase, on the other hand, is so efficient that it interferes with direct binding and kinetics studies, such that Klenow mutants with reduced or eliminated exonuclease activity have been used in almost all such studies for over a decade. Removal of the exonuclease catalytic activity does not, however, necessarily abolish DNA binding to the exonuclease site. In fact the co-crystal of Klenow bound to primer-template DNA utilizes one of these exonuclease deficient mutants, yet unexpectedly shows the DNA bound in a so called "editing mode" which includes contacts in the exonuclease domain, the edge of the polymerase domain cleft, and a second cleft that is formed upon DNA binding and runs roughly between the two

active sites and roughly perpendicular to the polymerase domain cleft (23). Co-crystal structures of DNA bound to Klentaq show the DNA bound further up into the polymerase active site cleft (although it does not pass all the way through the cleft) (24). The position of the DNA in the crystal structure of the binary complex of full length Taq and DNA (29) is similar to that for Klentaq. The binary complexes for Klenow and Klentaq are shown in Fig. 3.7. It may be that as exonuclease activity evolved, the strength of the interaction of the DNA with the exonuclease site shifted the location of the initial binding site for DNA. There are a billion years of evolutionary distance between the two bacterial families (30). This hypothesis of two different initial binding sites would be consistent with the available Klenow co-crystal structure (and the two-site shuttling model proposed for movement of the DNA between sites) (23), consistent with the calculations that were performed on the Klenow structural data (31) and consistent with the significantly higher binding affinity of Klenow for DNA versus Klentaq, and the larger linked ion release upon binding by Klenow relative to Klentaq. Further studies are required for elucidation of which, if any of these possibilities can explain the binding and salt sensitivity differences between the two species of polymerases.

$MgCl_2$ Dependence of DNA Binding by the Polymerases.

The $MgCl_2$ concentration ranges showing linked ion releases for Klentaq and Klenow are, like the KCl concentration ranges, separated by about an order of magnitude. Klentaq DNA binding releases 0.8 net Mg^{+2} ions in the 0-10 mM [$MgCl_2$] range, while Klenow DNA binding releases 1.1 net Mg^{+2} ions in the 10-50 mM [$MgCl_2$] range.

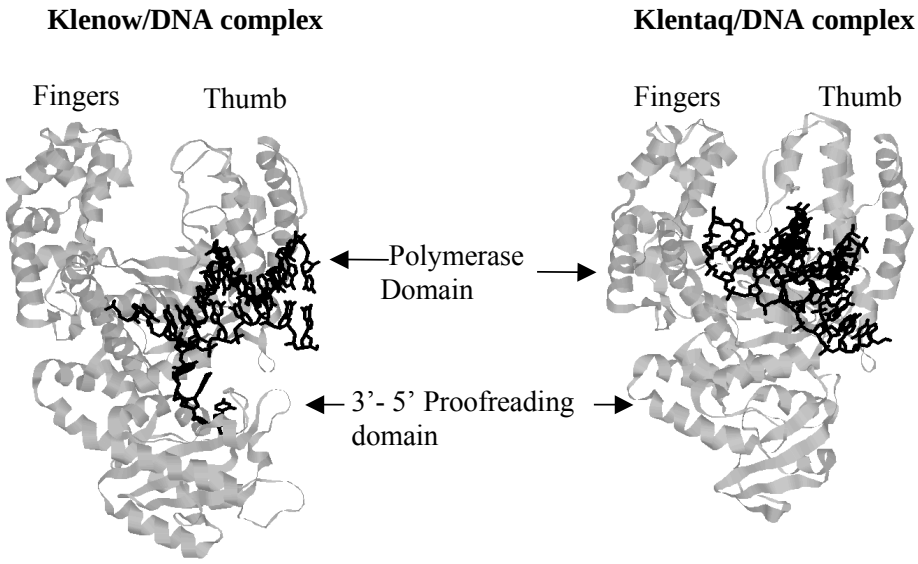


Figure 3.7. X-ray crystal structures of Klenow (Protein Data Bank (PDB) code 1KLN (23)) and Klentaq (PDB code 4KTQ (24)) polymerases bound to DNA.

As with the majority of DNA binding proteins assayed to date, the Mg^{+2} linkages are lower than the monovalent ion linkages. For many DNA binding proteins, the linked ion release in $MgCl_2$ is half that in monovalent salts, indicative of cation specific effects (13-15). Deviations from this ratio can be an indication of specific ion effects. The fact that the net Mg^{+2} releases for Klentaq and Klenow are lower than half the K^+ release is likely due to the fact that the $MgCl_2$ dependencies were (by necessity) determined in the presence of KCl (13-15). Because the net ion release is much lower for Mg^{+2} the statistical significance of the difference between the values for the two polymerases is less reliable. It is interesting to note, however, that similar to the findings with KCl, Klenow DNA binding is linked to the release of more Mg^{+2} than Klentaq DNA binding (about 30% more). Thus, like the results with KCl, the linked ion releases of Mg^{+2} upon DNA binding, and the significant shift in the salt ranges for similar affinity binding of the

two polymerases both suggest different binding conformations or binding footprints for the two polymerases in their respective protein-DNA complexes.

Binding of both species of polymerase to DNA is linked to the release of Mg^{+2} at higher $MgCl_2$ concentrations, but both species of polymerase appear to require divalent ions to achieve their highest DNA binding affinity. Treatment of the polymerases with EDTA lowers the DNA binding affinity of Klentaq and Klenow by ~6X ($\Delta\Delta G$ = 1.0 kcal/mole) and ~13X ($\Delta\Delta G$ = 1.53 kcal/mole), respectively. Removal of EDTA, however, almost completely restores the binding affinity originally observed in the absence of EDTA. This suggests that EDTA interacts with metals on the protein surface and either competitively (in the binding site) or allosterically diminishes DNA binding affinity, but does not remove the required metals from the protein at least at the concentration used in this experiment. Because EDTA also chelates other divalent cations, from this experiment alone it is not conclusive that the required ion is Mg^{+2}. Bound zinc has been reported for both species of polymerase (32,33). Both species of polymerase are known to require Mg^{+2} or Mn^{+2} for catalytic activity (34, 35, 36). *E. coli* Pol I has been reported to bind up to 21 magnesium and/or manganese ions using electron paramagnetic resonance (37).

Summary

We have examined the DNA binding properties of the type I DNA polymerases from *Escherichia coli* and *Thermus aquaticus* as a function of KCl and $MgCl_2$. The results of this study show that these two species of polymerase, when binding to the same DNA, 1) bind with very different affinities under similar solution and salt concentration conditions, 2) differ substantially in their sensitivity to the salt concentration range for binding, 3) release different numbers of K^+ and Mg^{+2} ions upon formation of the protein-

DNA complex, and 4) differ in their requirements for protein bound divalent cations. It is not yet determined which, if any, of these functional differences are correlated with the ability of Taq/Klentaq polymerase to function at such dramatically higher temperatures than Klenow.

References

1. Patel, P.H., Suzuki, M., Adman, E., Shinkai, A. and Loeb, L.A. (2001) *J. Mol. Biol*. **308**, 823-837.

2. Patel, P.H. and Loeb, L.A. (2001) *Nat. Struct. Biol*. **8,** 656-659.

3. Cowart, M., Gibson, K.J., Allen, D.J. and Benkovic, S.J. (1989) *Biochemistry* 28, 1975-1983.

4. Morales, J.C. and Kool, E.T. (2000) *Biochemistry* **39**, 12979-12988.

5. Barnes, W.M. (1992) *Gene (Amst.)* **112**, 29-35.

6. Derbyshire, V., Freemont, P.S., Sanderson, M.R., Beese, L., Friedman, J.M., Joyce, C.M. and Steitz, T.A. (1988) *Science* **240**, 199-201.

7. Bradford, M.M. (1976) *Anal. Biochem*. **72**, 248-254.

8. Kuchta, R.D., Mizrahi, V., Benkovic, P.A., Johnson, K.A. and Benkovic, S.J. (1987). *Biochemistry* **26**, 8410-8417.

9. Heyduk T. and Lee J.C. (1990) *Proc Natl Acad Sci U S A*. **87**, 1744-8.

10. Heyduk T., Ma, Y., Tang, H. and Ebright, R.H. (1996) *Methods Enzymol*. **274**, 492-503.

11. Inglese, J., Blatchly, R.A. and Benkovic, S.J. (1989) *J. Med. Chem*. **32**, 937-940.

12. Wyman, J. (1964) *Adv. Protein Chem*. **19**, 223-286.

13. Record, M.T., Zhang, W. and Anderson, C.F. (1998) *Adv. Protein Chem*. **51**, 281-353.

14. Record, M.T., Ha, J.H. and Fisher, M.A. (1991) *Methods Enzymol*. **208**, 291-343.

15. Lohman, T.M. and Mascotti, D.P. (1992) *Methods Enzymol.* **212**, 400-424.

16. Engelke D.R., Krikos, A., Bruck, M.E. and Ginsburg, D. (1990) *Anal. Biochem*. **191**, 396-400.

17. Brock, T.D. (1974) in *Bergy's Manual of Determinative Bacteriology*, 8th Edition, Eds.p. 285, Buchanan & Gibbons (Williams and Wilkins, Baltimore, MD).

18. Lawyer, F.C., Stoffel, S., Saiki, R.K., Chang, S.Y., Landre, P.A., Abramson, R.D. and Gelfand, D.H. (1993) *PCR Methods Appl.* **2**, 275-287.

19. deHaseth, P.L., Lohman, T.M. and Record, M.T. (1977) *Biochemistry* **16**, 4791-4796.

20. deHaseth, P.L., Lohman, T.M., Burgess, R.R. and Record, M.T. (1978) *Biochemistry* **17**, 1612-1622.

21. Lundback, T. and Hard, T., (1996) *J. Phys. Chem*. **100**, 17690-17695.

22. Heyduk, E., Baichoo, N. and Heyduk, T. (2001) *J. Biol. Chem*. **276**, 44598-44603.

23. Beese, L.S., Derbyshire, V. and Steitz, T.A. (1993) *Science*, **260**, 352-355.

24. Li, Y., Korolev, S. and Waksman, G. (1998) *EMBO J*. **17**, 7514-7525.

25. Barnes, W.M. (1995) (July 25, 1995) US Patent 5,436,149.

26. Denman, S., Hampson, K. and Patel, B.K.C. (1991) *FEMS Microbiol. Lett.* **82**, 73-78.

27. Madigan, M.T., Martinko, J.M., and Parker, J. (1997) in *Brock Biology of Microorganisms*, 8th Ed., p. 170, Prentice Hall, Upper Saddle River, NJ.

28. Barkley, M.D., Lewis, P.A. and Sullivan, G.E. (1981) *Biochemistry* **20**, 3842-3851.

29. Eom, S.H., Wang, J. and Steitz, T.A. (1996) *Nature* **382,** 278-281.

30. Battista, J. R. and F. A. Rainey. (2001) in *Bergey's Manual of Systematic Bacteriology* (Boone, D. R., and Castenholz, R.W., eds), 2nd Ed., pp. 395-414, Vol. 1, Springer-Verlag New York Inc., New York.

31. Yadav, P.N., Modak, M.J. and Yadav, J.S. (1994) *J. Mol Recognit*. **7**, 207-209.

32. Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G. and Steitz, T.A. (1985) *Nature* **313**, 762-766.

33. Kim, Y., Eom, S.H., Wang, J., Lee, D.S., Suh, S.W. and Steitz, T.A. (1995) *Nature* **376**, 612-616.

34. Chien, A., Edgar, D.B. and Trela, J.M. (1976) *J. Bacteriol.* **127**,1550-1557.

35. Englund, P.T., Huberman, J.A., Jovin, T.M. and Kornberg, A. (1969) *J. Biol. Chem*. **244**, 3038-3044.

36. Slater, J.P., Tamir, I., Loeb, L.A. and Mildvan, A.S. (1972) *J. Biol. Chem*. **247,** 6784-6794.

37. Mullen, G.P., Serpersu, E.H., Ferrin, L.J., Loeb, L.A. and Mildvan, A.S. (1990) *J. Biol. Chem.* **265**, 14327-14334.

CHAPTER 4

THERMODYNAMICS OF THE BINDING OF *THERMUS AQUATICUS* DNA POLYMERASE TO PRIMED-TEMPLATE DNA*

Introduction

In the previous chapter we examined the differences and similarities in the salt dependences of DNA binding by Taq and *E. coli* DNA polymerases (1). In this chapter, we have examined the temperature dependence of DNA binding by Taq polymerase and its Klentaq large fragment domain in order to begin to understand some of the thermodynamic driving forces and non-covalent interactions involved in the functioning of Taq polymerase. The energetic forces that drive a macromolecular interaction are the defining characteristics of that particular interaction. Understanding thermodynamic profiles for whole classes of interactions, such as DNA-protein interactions, allow one to begin to understand the energetic constraints or requirements for evolution and/or *de novo* design of such interactions. In this chapter, we have performed equilibrium DNA binding experiments with Taq polymerase using both fluorescence anisotropy and isothermal titration calorimetry, and determined, as a function of temperature, the core thermodynamic quantities for this interaction (ΔG, ΔH, ΔS, ΔCp).

We find that DNA binding for Taq and Klentaq occur with sub-micromolar affinity across a broad temperature range, with maximal affinities near 40-50^{o}C. Although it has been shown that Taq polymerase has little or no catalytic activity at room temperature (2,3), we find that Taq and Klentaq polymerases bind DNA quite well down

---

* Reprinted with permission from *Nucleic Acids Research*. **Vol. 31(19),** 5590-97.

to at least 5°C. We also find that the DNA binding of Taq/Klentaq is associated with an unusually large heat capacity change for a non-sequence specific binding protein (4-8). Because of this, as found for most sequence-specific DNA binding proteins (8), the driving force for DNA binding of Taq/Klentaq shifts from entropy driven to enthalpy driven as the temperature is increased. At its physiological temperature, DNA binding is enthalpy driven for Taq/Klentaq polymerase.

Materials and Methods

Materials

The proteins examined are full length Taq DNA polymerase and the Klentaq large fragment of the polymerase. Preparation of the proteins has been described in detail previously (9). No surfactants are used during preparation, storage, or experiments with the polymerases. Fluorescently labeled and unlabeled DNA oligodeoxyribonucleotides were purchased from IDT (Integrated DNA Technologies Inc).

Fluorescence Anisotropy Assay

Equilibrium DNA binding experiments were performed with the following primer-template set:

63/70mer:

```
5′ TACGCAGCGTACATGCTCGTGACTGGGATAACCGTGCCGTTTGCCGACTTTCGCAGCCGTCCA3′
3′ ATGCGTCGCATGTACGAGCACTGACCCTATTGGCACGGCAAACGGCTGAAAGCGTCGGCAGGTTCCCAAA5′
```

For fluorescence anisotropy assays, the primer was labeled at the 5'end with Rhodamine-X (ROX). This DNA has been used previously for salt dependence studies of Taq DNA binding (Chapter 3 and Ref 9), and contains the same DNA sequence used previously for kinetic studies of DNA binding of Klenow polymerase by Benkovic and associates (10) at its single-strand double-strand junction. The extended length of this DNA allows it to remain stable to higher temperatures. Titrations were performed in

10mM Tris, 75mM KCl, 5mM $MgCl_2$, pH 7.9 buffer at the indicated temperatures. pH of the buffers were adjusted at each experimental temperature.

The pH was adjusted by mixing Tris base and Tris-HCl. ROX-labeled DNA was titrated with increasing concentrations of protein and binding was monitored using the anisotropy signal change as protein-DNA complex is formed (9). Fluorescence anisotropy measurements were performed using a FluoroMax-2 fluorometer equipped with an automated polarizer and regulated at the indicated temperatures. The excitation and emission wavelengths were 583nm and 605nm respectively, with 8nm band-pass and integration time of 10 seconds. For all the equilibrium titrations the DNA concentration used was 1nM. In all the experiments the protein was titrated into fluorescently labeled DNA, with the total [DNA] < K_d. After each addition the sample was equilibrated at the required temperature for 8 min and anisotropy was measured. Temperature was varied across the widest possible range for each polymerase. Low temperature titrations were performed down to 5^{o}C. Limiting values at high temperatures were the point where well-behaved, reproducible isotherms could no longer be obtained. Titrations at high temperatures were performed with a screw cap cuvette to minimize evaporation during the experiments, and at low temperatures nitrogen was flowed through the sample chamber to prevent condensation on the outside of the cuvette.

Isothermal Titration Calorimetry

Titrations were performed as a function of temperature on a MicroCal VP-ITC for the binding of Taq and Klentaq to the same 63/70mer DNA described above, without the fluorescent label. Titrations were performed in the same buffers as the fluorescence titrations. Taq or Klentaq were titrated into 1μM DNA. Each titration consisted of an initial 2μl injection (not used for data analysis) followed by 31-36 subsequent 7μl

injections, except for the 50^{o} and 60^{o} titrations for full length Taq, which consisted of 24 subsequent 10μl injections. The heat of dilution of the protein was obtained by titrating protein into the buffer. The actual heat of the reaction was determined after subtracting the heat of dilution of the protein. ITC binding curves were analyzed using the single-site binding equation in the MicroCal Origin software package.

Data Analysis

Equilibrium binding curves obtained using fluorescence anisotropy were fit to a standard single-site isotherm usable when the [DNA] << K_d:

$$\Delta A = [\Delta A_T (E_T/K_d)/(1+ E_T/K_d)] \quad \textbf{Eq. 4.1}$$

where ΔA is the change in fluorescence anisotropy, ΔA_T is the total change in anisotropy, E_T is the total polymerase concentration at each point in the titration, and K_d is the dissociation constant for polymerase-DNA binding. It was shown previously that the polymerases bind the DNA with 1:1 stoichiometry (Chapter 3 and Ref. 9). Titrations with low K_d values were also checked with a more general equation which does not assume equality between the total and free enzyme concentrations, as described previously (9). Fits to the titrations with the very tightest (lowest) K_d values give ≤ 10% variation in the fitted K_d values using the two different equations. All other titrations give, within error, identical fitted K_d values with Equation 1 and the more general equation (9). All non-linear fitting was performed using the program KaleidaGraph (Synergy Software). Temperature dependencies of equilibrium binding were analyzed using an integrated form of the Gibbs-Helmholtz equation:

$$\Delta G_{(T)} = \Delta H_{refT} - T\Delta S_{refT} + \Delta Cp\,[T - T_{refT} - T\ln(T/T_{refT})] \quad \textbf{Eq. 4.2}$$

where $\Delta G_{(T)}$ is the free energy at each temperature (dependent variable), T is the temperature in Kelvin (independent variable), ΔCp is the heat capacity, and ΔH_{refT} and ΔS_{refT} are the fitted van't Hoff enthalpy and entropy values at any chosen "reference temperature" T_{refT}.

Circular Dichroism (CD) Measurements

CD spectra were measured at room temperature (22°C) in an AVIV Model 202 CD spectrophotometer. A dual compartment mixing cuvette (Starna Cells), was used to record the spectra of protein + DNA before and after mixing. One compartment was filled with 3 μM DNA, the other with 3μM Klentaq.

Results

Direct equilibrium binding of Taq and Klentaq polymerases to 63/70mer DNA was measured using a fluorescence anisotropy assay over a temperature range of 5-70°C. Selected titration curves are shown in Figure 4.1. Each titration curve fits well to a single- site binding isotherm, and it can be seen from the precision of the data that even modest shifts in K_d can be readily quantitated. The binding affinity increases with temperature until about 40-50°C, after which binding affinity decreases with increasing temperature. The titrations shown in Figure 4.1 illustrate the behavior of the titration curves in these increasing and decreasing portions of temperature space. In Figure 4.2, the temperature dependence of the binding equilibrium is shown as a Gibbs-Helmholtz plot (ΔG versus T) for binding of Taq and Klentaq polymerases to DNA. The observed non-linearity of ΔG versus T results from a negative heat capacity change (ΔCp) associated with the binding process, and ΔCp can be calculated using the Gibbs-Helmholtz equation, which allows direct calculation of the van't Hoff enthalpy (ΔH_{vH})

and entropy (ΔS_{vH}) as a function of temperature. The binding parameters at each temperature are listed in Table 4.1.

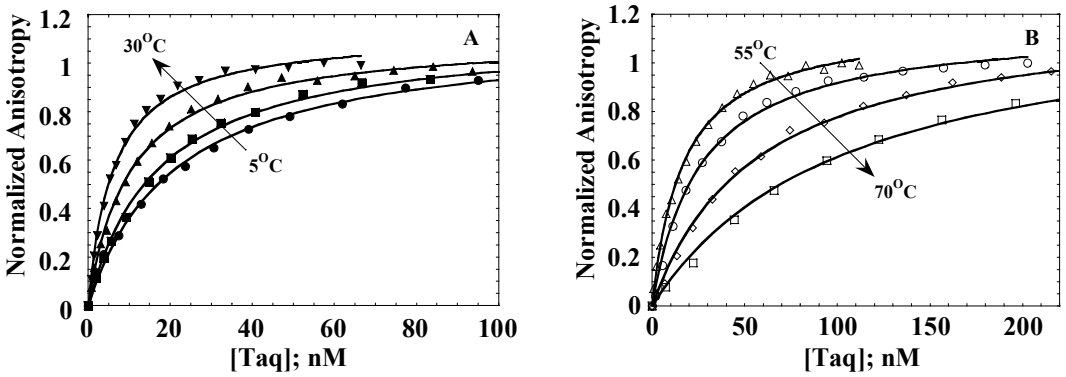


Figure 4.1. Equilibrium titrations of the binding of Taq polymerase to DNA monitored using fluorescence anisotropy. A) Representative titrations of Taq-DNA binding at lower temperatures 5°C(○), 10°C (□), 20°C (▲), 30°C (▼). B) Representative titrations of Taq-DNA binding at higher temperatures, 55°C (Δ), 60°C (○), 65°C (◊), 70°C (□).

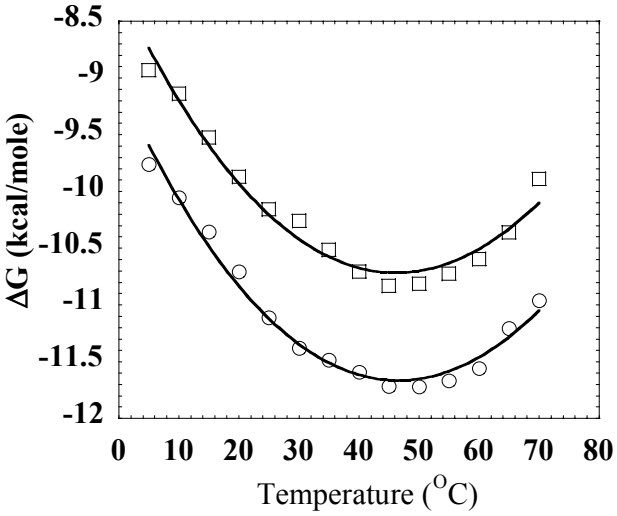


Figure 4.2. Gibbs-Helmholtz plot showing the temperature dependence of the free energy (ΔG) of DNA binding for Taq (○) and Klentaq (□). Lines are the fits to the Gibbs-Helmholtz equation (Eq. 4.2) as described in Materials and Methods.

Table 4.1. Thermodynamic Parameters of DNA Binding by Taq and Klentaq DNA Polymerases[a]

Taq

Temp (°C)	K_d (nM)	ΔG (kcal/mol)	ΔH (kcal/mol)	ΔS (kcal/mol K)	TΔS (kcal/mol)	ΔCp (kcal/mol K)
5	21.3 ± 1.0	-9.76	+18.81	+0.102	+28.38	
10	17.3 ± 0.6	-10.05	+15.14	+0.089	+25.22	
15	13.8 ± 0.5	-10.36	+11.47	+0.076	+21.94	-0.73±0.04
20	10.3 ± 0.4	-10.71	+7.81	+0.064	+18.63	
25	7.1 ± 0.2	-11.11	+4.14	+0.051	+15.26	
30	6.2 ± 0.3	-11.38	+0.47	+0.039	+11.81	
35	7.1± 0.2	-11.49	-3.2	+0.027	+8.31	
40	8.1± 0.3	-11.59	-6.86	+0.015	+4.75	
45	8.9 ± 0.4	-11.71	-10.53	+0.004	+1.11	
50	11.7 ± 0.6	-11.72	-14.2	-0.008	-2.54	
55	16.8 ± 0.6	-11.67	-17.86	-0.019	-6.26	
60	26.0 ± 1.4	-11.56	-21.53	-0.030	-10.07	
65	56.7 ± 2.4	-11.21	-25.2	-0.041	-13.92	
70	103.7 ± 6.5	-10.96	-28.87	-0.052	-17.82	

Klentaq

Temp (°C)	K_d (nM)	ΔG (kcal/mol)	ΔH (kcal/mol)	ΔS (kcal/mol K)	TΔS (kcal/mol)	ΔCp (kcal/mol K)
5	95.0 ± 1.7	-8.93	+18.62	+0.098	+27.36	
10	87.7± 1.8	-9.14	+15.06	+0.086	+24.25	
15	59.0 ± 0.8	-9.53	+11.5	+0.073	+21.08	-0.71± 0.06
20	43.3± 1.1	-9.87	+7.93	+0.061	+17.86	

(table cont'd)

25	35.6 ± 1.1	-10.16	+4.37	+0.049	+14.57
30	39.8 ± 1.4	-10.26	+0.81	+0.037	+11.23
35	34.6 ± 1.1	-10.51	-2.75	+0.025	+7.82
40	33.3 ± 1.3	-10.71	-6.32	+0.014	+4.36
45	35.8 ± 1.5	-10.83	-9.88	+0.003	+0.83
50	48.3 ± 1.5	-10.81	-13.44	-0.008	-2.74
55	71.4 ± 3.0	-10.72	-17.0	-0.019	-6.36
60	111.4 ± 5.7	-10.59	-20.57	-0.030	-10.06
65	199.7 ± 9.1	-10.36	-24.13	-0.041	-13.8
70	497.5 ± 23.3	-9.89	-27.7	-0.051	-17.59

[a]K_d and ΔG values are the experimental values determined from titrations at each temperature. The other thermodynamic parameters are calculated from the fit to the Gibbs-Helmholtz equation. Errors are the parameter value errors returned from the fits to Eq. 4.1 (K_d values) or Eq. 4.2 (ΔCp, ΔH and ΔS values).

Figure 4.3 shows thermodynamic profiles for DNA binding of Taq and Klentaq polymerases, including the values of ΔG, ΔH_{vH}, and $T\Delta S_{vH}$, as a function of temperature. Plotted on this scale (about 20X the scale in Figure 4.2) it can be seen that the magnitudes of the temperature dependent enthalpy and entropy changes are far larger than the temperature deviation of ΔG. In other words, the ΔH_{vH}, and $T\Delta S_{vH}$ are strongly temperature-dependent. The binding of the polymerases exhibits enthalpy-entropy compensation: where the ΔH and $T\Delta S$ change in parallel with temperature (11,12). This thermodynamic profile is generally characteristic of sequence-specific DNA binding proteins (8,11,12).

The enthalpy of binding for Taq and Klentaq to the 63/70mer DNA was also measured calorimetrically. Titrations were performed under stoichiometric conditions to

measure the heat of the reaction at several different temperatures, and representative titrations at high and low temperature are shown in Figure 4.4. The buffer conditions

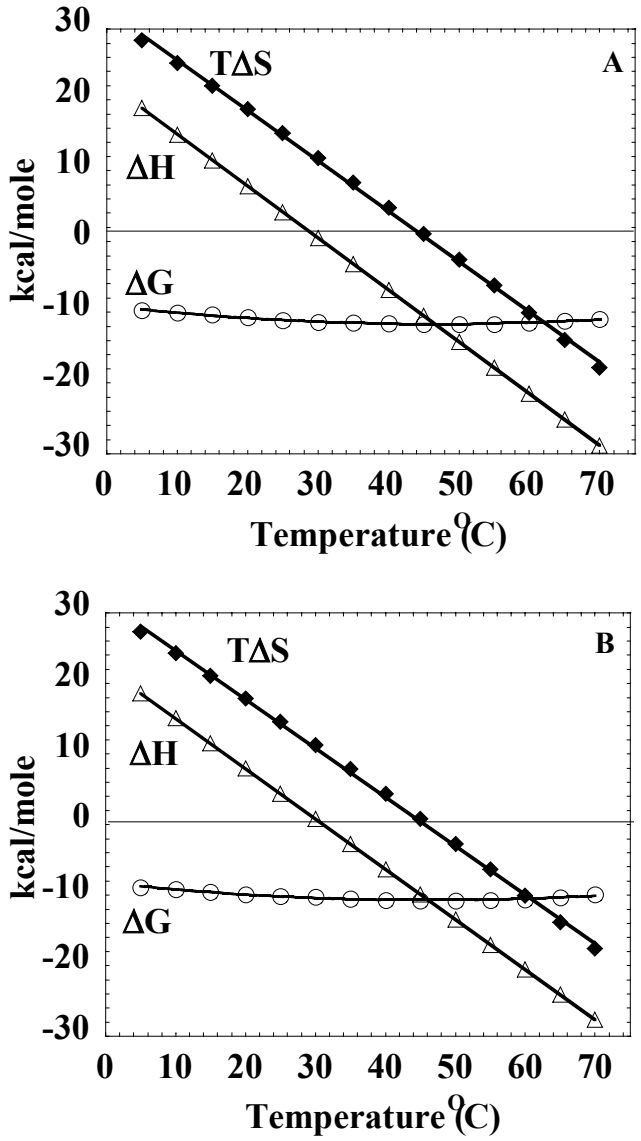


Figure 4.3. Temperature dependencies of the thermodynamic parameters ΔH (Δ), TΔS (♦) and ΔG (○) of DNA binding by A) Taq and B) Klentaq. ΔG is the same data as in Figure 4.2, plotted here on a different scale to show its relationship to the ΔH and TΔS parameters. ΔH and ΔS values at each temperature are derived from the non-linear analysis of the Gibbs-Helmholtz plot.

were identical to those used to determine the temperature dependence of equilibrium binding. There are limitations in performing calorimetric titrations that are not a problem with the anisotropy measurements. For example, calorimetric titrations require 50-100 fold more protein and DNA per titration than do the fluorescence titrations. In addition, the binding enthalpy for Taq/Klentaq changes from positive to negative and thus passes through zero, so there is a temperature span near the middle of the binding range where the enthalpy change is unmeasureable calorimetrically using any reasonable quantities of protein and DNA. Notwithstanding, we calorimetrically determined the DNA binding enthalpies for Taq and Klentaq at several temperatures between 10 and 60^{o}C. Data are reported in Table 4.2, and are shown in Figure 4.5 along with the ΔH_{vH} dependence from Gibbs-Helmholtz analysis (from Figure 4.2).

While there is some variability in the absolute ΔH's returned by van't Hoff versus calorimetric analysis at any one temperature, the temperature dependencies of the calorimetric and van't Hoff binding enthalpies (and the ΔCp values calculated from them) agree quite well. This agreement is an important confirmation of the thermodynamics, as differences between calorimetric and van't Hoff enthalpies and heat capacities are observed in a large fraction of the systems that have been examined with both approaches and the origins of these common discrepancies has been the subject of ongoing debate for nearly a decade [see Liu and Sturtevant (13) and Horn et. al. (14) for two of the most recent studies]. The temperatures T_H (where ΔH is 0) and T_S (where ΔS is 0) are ~32^{o}C and ~45^{o}C respectively for Taq and Klentaq. The T_H represents the temperatures at which K_d is minimum and T_S represents the temperature at which ΔG is minimum (15).

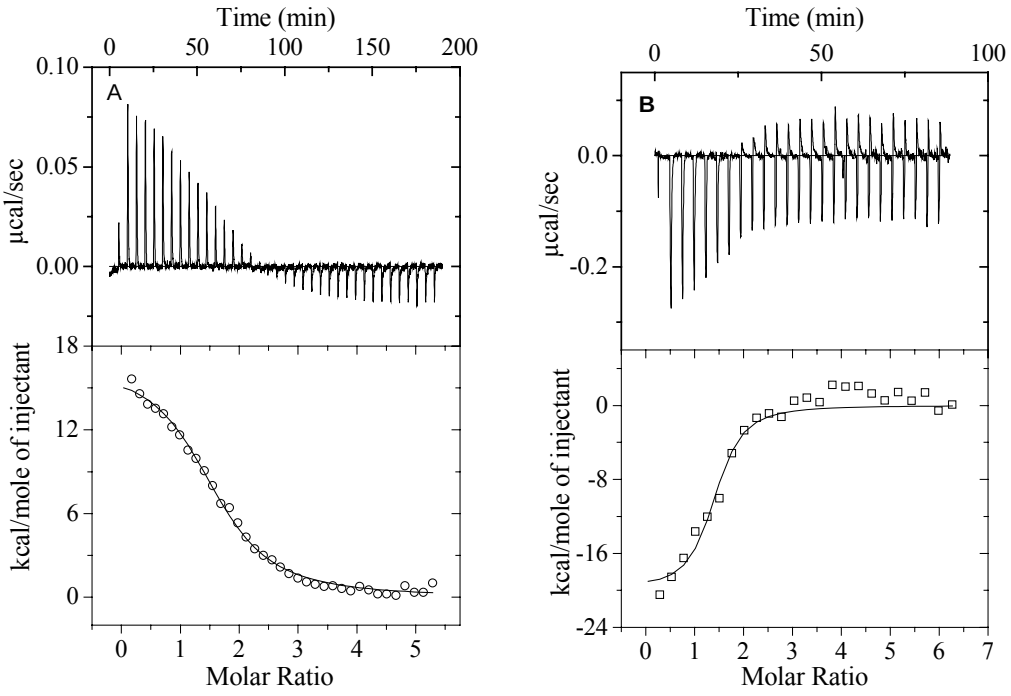


Figure 4.4. Calorimetric titrations of Taq into DNA at A)10°C and B) 60°C. Shown here are the raw calorimetric data (upper panel of each pair) and the fits to the integrated heats of the reaction (lower panel of each pair).

Table 4.2. Calorimetric ΔH and ΔCp Values for Taq and Klentaq

	Taq		Klentaq	
Temperature(°C)	**ΔH_{cal} (kcal/mole)**	**ΔCp_{cal} (kcal/mol K)**	**ΔH_{cal} (kcal/mole)**	**ΔCp_{cal} (kcal/mol K)**
10	+16.73		+20.93	
20	+15.21	-0.80 ± 0.09	+12.75	-0.82 ± 0.04
50	-14.64		-8.93	
60	-19.95		-21.55	

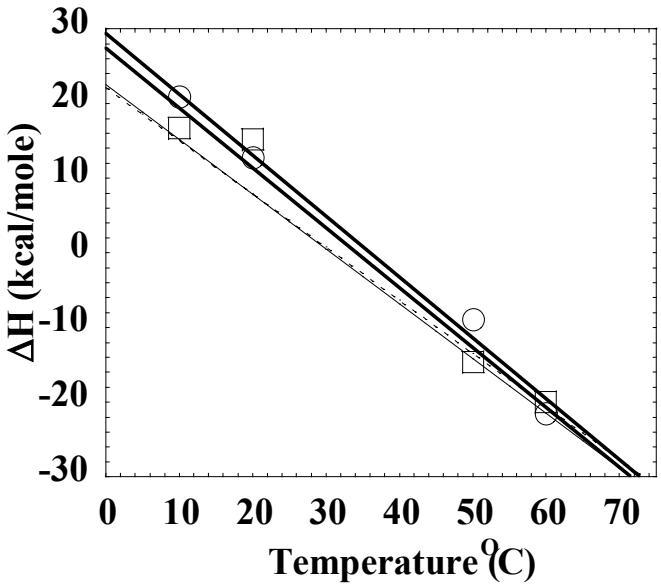


Figure 4.5. Temperature dependence of the enthalpy change (ΔH) upon binding of Taq and Klentaq to DNA. Calorimetric enthalpy (ΔH_{cal}) values obtained from the calorimetric titrations for Taq (□) and Klentaq(○) are shown along with linear fits to the calorimetric data used to obtain the calorimetric ΔCp (bold lines). For comparison, the van't Hoff enthalpy temperature dependencies derived from the Gibbs-Helmholtz analysis is also shown here as thin solid line for Taq and as thin dashed line for Klentaq (these two lines almost overlap).

To examine if DNA binding of the polymerase is associated with changes in protein or DNA structure in solution, we examined the circular dichroism (CD) spectra of the Klentaq-DNA complex relative to the isolated protein and DNA. Experiments were carried out in a dual compartment mixing cuvette, as described in Materials and Methods, to insure that small spectral changes are exclusively due to complex formation. Data are shown in Figure 4.6 for the combined DNA + protein spectra before and after formation of the complex. Spectral signals above 240 nm in CD spectroscopy will be almost exclusively due to the DNA, while those below 240 nm are largely due to the protein. Figure 4.6 shows that there are small but easily observed conformational changes in both the DNA and the protein upon formation of the complex.

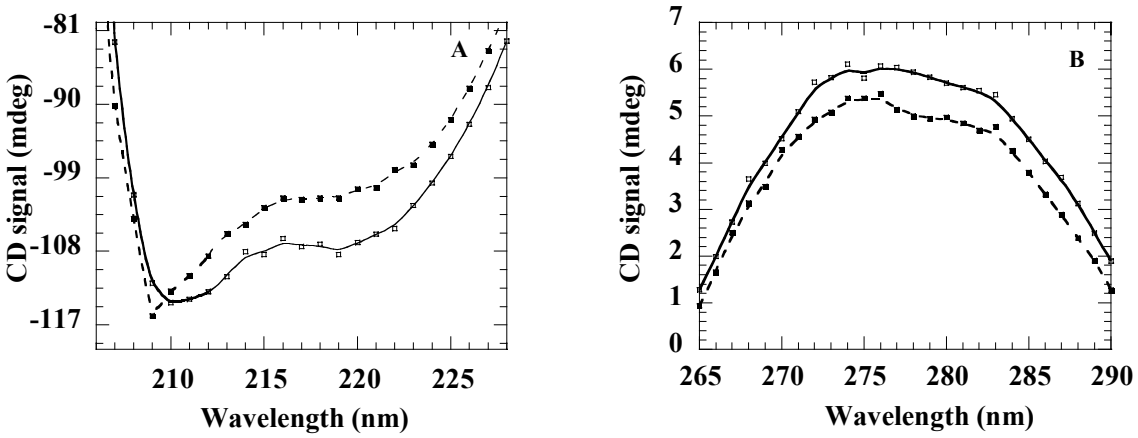


Figure 4.6. Circular dichroism spectra for the Klentaq-DNA complex (■) versus the free DNA + free protein (□) at 22°C. Spectra were obtained using a dual compartment mixing cuvette and recorded before and after mixing. Shown here are the regions of the spectra showing differences. A) the spectral region from 206 –228nm corresponding to signal primarily from the protein. B) the spectral region from 265-290nm corresponding to signal primarily from the DNA. Spectra were collected in 1nm steps and the lines are simply a smoothed interpolation of the stepwise spectral data.

Discussion

We have examined the temperature dependence of the thermodynamics of binding of Taq and Klentaq polymerases to DNA. Equilibrium binding occurs with dissociation constants (K_d) in the sub-micromolar range across a wide temperature span. The temperature dependence of the binding free energy (Gibbs-Helmholtz plot) is strongly nonlinear, indicating a temperature dependent binding enthalpy (ΔH_{vH}) and thus an associated ΔCp of binding. For comparison, absence of a ΔCp of binding would yield a linear dependence of ΔG on temperature. Parallel determination of the enthalpy of binding versus temperature by isothermal titration calorimetry confirms that the binding of Taq and Klentaq to DNA are associated with ΔCp values of –0.7 to –0.8 kcal/mole K.

These are substantial ΔCp values, and are in contrast to the absence of ΔCp generally associated with non-specific DNA-protein interactions (4-8).

Very few other thermophilic DNA binding proteins have had their DNA binding activity thermodynamically characterized. These include the non-specific binding of Sac7d from *Sulfolobus acidocaldarius* (7) and Sso7d from *Sulfolobus solfataricus* (16), and the site-specific binding of ORF56 protein from *Sulfolobus islandicus* (17). The site-specific binding of ORF56 has a ΔCp of –1.5 kcal/mole K (17). The non-specific binding of Sso7d and Sac7d have ΔCp values of –0.26 kcal/mole K and zero, respectively (7,16). The finding that the optimal DNA binding temperature is significantly below the optimal growth temperature for the thermophilic bacterium was also found for the DNA binding of ORF56 (17). An optimal binding temperature could not be determined for Sso7d, and Sac7d has no ΔCp of binding and hence has no minimum in its Gibbs-Helmholtz distribution (7,16).

Taq polymerase replicates DNA essentially non-sequence specifically, beginning at any single-stranded/double-stranded junction, and this non-specificity is a basis for Taq's utility in the polymerase chain reaction. While DNA structural elements such as mismatched base pairs, overhanging primers, gaps and nicks, et cetera, are known to have significant influence over DNA polymerase binding and function, the DNA sequence dependence of the binding of prokaryotic DNA polymerases to matched primed DNA has not been studied in any detail. They are certainly not sequence specific in the same way that transcriptional regulators or restriction endonucleases are specific. Preferences for different sequences have been found for several non-specific DNA binding proteins (6,7), including Klenow polymerase (10), but the difference between the tightest and weakest binding sequences is up to about an order of magnitude, in contrast to the 3 to 7 orders of

magnitude differences between the specific and non-specific binding modes of site-specific DNA binding proteins (8). If any binding sequence specificity does exist for DNA polymerases, it is possible that the thermodynamic profile for DNA binding of a polymerase could be different for alternate DNA sequences.

Is a measurable ΔCp of binding a common characteristic of proteins that bind DNA primarily non-specifically or that bind to specific DNA structures? Non-specific binding of DNA is generally correlated with a near zero ΔCp (reviewed in 8). Correspondingly, a negative ΔCp has been called a "signature" of site-specific DNA binding (8). However, a few exceptions to this correlation have been reported, namely the ΔCp values reported for non-specific DNA binding of *E. coli* SSB and the Sso7d protein of *S. solfataricus* and *E.coli* IHF (6,7,18). Our study of the binding of Taq/Klentaq to DNA is the first characterization of ΔCp (along with ΔH and ΔS and their temperature dependencies) for the DNA binding of a DNA polymerase.

The non-specific protein-DNA interactions thus far found to have near zero ΔCp values are mostly sequence-specific binding proteins binding in their non-specific mode (Sac7d from *S. solfataricus* is an exception) (8, 16). The three non-specific binding proteins that have been found to bind DNA with significant ΔCp values (Taq, SSB, and Sso7d) differ from other non-specific binding proteins in two fundamental ways. Taq, SSB, and Sso7d are proteins that primarily bind non-sequence specifically to DNA (although with some sequence preferences), and are proteins that show DNA structure specificity or preferences. These distinctions may be indicative of different subsets of non-specific DNA binding with different thermodynamic profiles. It may be that structure specific DNA binding, like sequence-specific binding, will be found to be

generally correlated with a larger ΔCp. As with sequence-specific binding, the formation of a tight complementary interface in a structure-specific DNA complex might be expected. On the other hand, as suggested previously in the case of Sso7d, a measurable ΔCp may be a general characteristic of proteins that primarily bind non-specifically to DNA (7). However, until more primarily non-sequence dependent binding proteins are examined, it is unclear how many thermodynamic/functional subclasses of such proteins might exist.

Conformational Changes Associated with Complex Formation

The circular dichroism (CD) spectrum of the Klentaq-DNA complex when compared to the spectrum of the combined (but physically separated) DNA + polymerase clearly shows that the conformations of both the DNA and the protein change slightly upon complex formation. The CD spectrum at wavelengths above about 240 nm is due primarily to the DNA, and the 265-290nm spectrum in Figure 4.6 is characteristic of B form DNA. CD spectral changes have been used to detect protein and DNA conformational changes upon binding (16,19), and a lack of spectral changes in this wavelength range has been used to demonstrate the absence of DNA conformational changes upon binding (20). The CD spectral changes shown in Figure 4.6, while small, demonstrate that the association of Klentaq with DNA is not a simple rigid body association. Conformational changes of the protein upon DNA binding are also observed by crystallography and are discussed further below.

Possible Molecular Bases for the Negative ΔCp of DNA Binding of Taq

The data in this study do not allow us to pinpoint the molecular contributions to the ΔCp of binding of Taq to DNA, but we can discuss some aspects of the possible correlations with the major proposed molecular models for ΔCp effects. For many

protein-DNA interactions and protein folding reactions there is a correlation between ΔCp and surface area changes (ΔASA) (21-26). This has proven a powerful connection between thermodynamic and structural information. However, there have also been a number of exceptions to this correlation (8,18,27-32) suggesting that such correlations may represent a specific subset of all reactions with a ΔCp. The exceptions often have the problem that the ΔCp for the process of interest is much larger than can be accounted for by burial of surface area. This is not the case for Taq/Klentaq. A ΔCp of –0.765 (mean of the Gibbs-Helmholtz and calorimetric ΔCp's), while unusually high for a non-sequence specific protein-DNA interaction, is low compared to many sequence-specific binding proteins. Structure based calculations of the surface area buried in the binding interface between Taq and DNA yield a value of 2530 $Å^2$ (33). If the entire change in surface area for the protein and DNA are included (i.e. not just the interface region), this total increases to 3126 $Å^2$ (LiCata, unpublished data). At least four different quantitative relationships between ΔCp and the sum of buried non-polar + polar surface area have been proposed (22-25). A correlation between ΔCp and buried surface area for Taq/Klentaq can easily be achieved using any one of these relationships. However, all of these equations predict that a relatively high fraction of the buried surface area must be non-polar: the mean of all four relationships predicts that 82% of the buried surface area must be non-polar. Since protein-DNA interfaces are generally much more polar than interiors of proteins, it would seem that other linked processes may also be involved in the generation of the ΔCp of binding of Taq to DNA.

Another major proposed origin of negative ΔCp values for protein-DNA interactions are conformational changes of the DNA upon binding (30,34). Binding of

TBP to the E4 promoter has been shown to be associated with large negative ΔCp that could not be explained by any significant burial of non-polar surfaces or conformational changes of the protein (30). It was proposed that the negative ΔCp might originate from the unwinding of the B-DNA helix, base unstacking, and intercalation of phenylalanine side chains in the DNA kinks (30). A large negative ΔCp has also been reported for the interaction of *E. coli* SSB with single-stranded DNA (6). Subsequent studies showed that this ΔCp is partly due to adenine base unstacking and partly due to a linked protonation equilibrium (34,35). The CD spectra for Klentaq binding to DNA in Figure 6 clearly show that there is some distortion of the DNA upon binding in the Klentaq-DNA complex. The crystal structures of DNA bound to Klenow and Klentaq (36-38) also show distortion of the DNA upon binding. Thus, structural changes in the DNA are also a likely contributor to the measured ΔCp of binding of Taq to DNA.

Low Temperature DNA Binding by Taq Polymerase

Surprisingly, Taq/Klentaq binds DNA with high affinity at temperatures as low as 5^{o}C. This was unexpected since it has been shown that Taq is essentially catalytically inactive at room temperature (2,3). The enzyme is optimally catalytically active at 70-75^{o}C (2,3). This indicates that catalysis but not binding involves a molecular process that can only occur at higher temperatures. The combination of high temperature stability with low temperature activity is rare in natural thermophilic proteins, and all but a very few thermophilic proteins are inactive or almost inactive at room temperature (reviewed in 39). This combination of properties can, however, be introduced into proteins by engineering or directed evolution (39).

In several studies, thermophilic proteins have been found to exhibit an increased molecular rigidity at room temperature, and this rigidity has been correlated with a

decreased ability to undergo the conformational fluctuations required for catalytic activity (40,41). Exceptions to this general correlation have also been observed (42). Although no data yet exist regarding the molecular rigidity of Taq, crystal structures have recently been solved for the binary complex of Klentaq polymerase with DNA, and the ternary complex of Klentaq with DNA and an incoming nucleotide (38). The binary complex structure corresponds to the interaction formed in our binding studies, the ternary complex reflects the enzyme performing catalytic polymerization. What is notable is that the structures of both complexes show significant conformational changes relative to the unbound polymerase (38). When Klentaq binds DNA, the "thumb" region rotates as a whole and moves in closer to the DNA (38). When the incoming nucleotide is added, the "fingers" region of the polymerase on the opposite side of the DNA binding cleft performs a similar closure motion (38). The change associated with catalysis is somewhat larger in magnitude, based on the number of atoms involved and the distances they move, but the structural change associated solely with DNA binding is itself a substantial one. The CD spectral changes upon DNA binding (Figure 4.6) confirm that in solution conformational changes occur in both the protein and the DNA. So, even though conformational flexibility is involved in the DNA binding process for Taq, it proceeds without a problem at low temperature.

The fact that Taq undergoes a conformational change upon DNA binding does not mean that it is not more rigid than mesophilic polymerases in general. In fact, the overall binding affinity of Taq for DNA was recently shown to be ~150X weaker than the affinity of Klenow polymerase for the same DNA under comparable solution conditions (9). This may be indicative of an overall higher rigidity of Taq polymerase relative to Klenow. Unlike the situation with the catalytic activity of many other thermophilic-

mesophilic protein pairs, however, there is no "corresponding states" temperature where the DNA affinity of Taq polymerase approximates that of Klenow polymerase.

While there have been numerous studies of the temperature dependencies of the enzymatic activity of thermophilic proteins, there have been very few direct substrate binding or substrate analog binding studies. It is interesting to note that in addition to Taq, the other thermophilic DNA binding proteins that have been characterized as function of temperature, Sac7d, Sso7d, and ORF56 from the *Sulfolobus* genus, which were discussed above, also bind DNA at temperatures below room temperature – although the unusual nature of this apparently common characteristic has not been noted previously. The DNA binding of Sac7d, Sso7d, and ORF56 were studied from 10-40^{o}C, 15 to 45^{o}C, and 17 to 57^{o}C, respectively (7,16,17). Thermophilic archaebacterial flap endonucleases have also previously been shown to bind DNA at low temperatures (43).Thus it would seem based on the examples available thus far, that unlike the usual loss of catalytic activity, the DNA binding activity of thermophilic proteins is maintained at lower temperatures.

Concluding Summary

The temperature dependence of DNA binding by Taq polymerase provides the first such characterization of DNA binding by a thermophilic protein that is not from the archaeon *Sulfolobus*. The two most unusual features of the thermodynamics of Taq-DNA interactions are the low temperature DNA binding of Taq, and the fact that its ΔCp and thermodynamic temperature profile are characteristic of sequence-specific DNA binding proteins. We have suggested that it is not Taq itself which is an exception to current empirical correlations, but that these correlations may need further refinement as our knowledge base on thermophilic and DNA binding proteins continues to expand.

References

1. Brock, T.D. (1974) in *Bergey's Manual of Determinative Bacteriology*, 8th Edition, Editors: Buchanan & Gibbons, Williams and Wilkins, Baltimore, MD, p. 285.

2. Lawyer, F.C., Stoffel, S., Saiki, R.K., Chang, S.Y., Landre, P.A., Abramson, R.D. and Gelfand, D.H. (1993) *PCR Methods Appl.* **2**, 275-287.

3. Hogrefe, H.H., Cline, J., Lovejoy, A.E. and Nielson, K.B. (2001) *Methods Enzymol*. **334**, 91-116.

4. Takeda, Y., Ross, P.D. and Mudd, C.P. (1992) *Proc Natl Acad Sci U S A* **89**, 8180-8184.

5. Ladbury, J.E., Wright, J.G., Sturtevant, J.M. and Sigler, P.B. (1994) *J. Mol. Biol.* **238**, 669-681.

6. Ferrari, M. E. and Lohman, T. M. (1994) *Biochemistry*. **33**, 12896-12910.

7. Lundback, T., Hansson, H., Knapp, S., Ladenstein, R. and Hard, T. (1998) *J. Mol.Biol*. **276**, 775-786.

8. Jen-Jacobson, L., Engler, L.E., Ames, J. T., Kurpiewski, M.R. and Grigorescu, A. (2000) *Supramolecular Chemistry* **12**, 143-160.

9. Datta, K. and LiCata, V.J. (2003) *J. Biol. Chem*. **278**, 5694-5701.

10. Kuchta, R.D., Mizrahi, V., Benkovic, P. A., Johnson, K. A. and Benkovic, S. J. (1987) *Biochemistry*. **26**, 8410-8417.

11. Ha, J.H., Spolar, R.S. and Record, M.T. (1989) *J. Mol. Biol*. **209**, 801-816.

12. Record, M.T., Ha, J.H., and Fisher, M.A. (1991) *Methods Enzymol* **208**, 291-343.

13. Liu, Y.F. and Sturtevant, J.M. (1997) *Biophys. Chem*. **64**, 121-126.

14. Horn, J.R., Brandts, J.F. and Murphy, K.P. (2002) *Biochemistry* **41**, 7501-7507.

15. Spolar, R.S., Ha, J.H. and Record Jr., M.T. (1989) *Proc. Natl. Acad. Sci. USA.* **86,** 8382-8385.

16. McAfee, J.G., Edmondson, S.P., Zegar, I. and Shriver, J.W. (1996) *Biochem*. **35**, 4034-4045.

17. Lipps, G., Stegert, M. and Krauss, G. (2001) *Nuclec Acids Research*, **29**, 904-913.

18. Holbrook, J.A., Tsodikov, O.V., Saecker, R.M. and Record, M.T., Jr (2001) *J. Mol. Biol.* **310**, 379-401.

19. Woody, RW. (1995) *Methods Enzymol*. **246**, 34-71.

20. Heyduk, E., Baichoo, N. and Heyduk, T. (2001) *J. Biol. Chem*. **276**, 44598-44603.

21. Sturtevant, J.M. (1977) *Proc. Natl. Acad. Sci. USA.* **74**, 2236-2240.

22. Spolar, R.S., Livingstone, J.R. and Record, M.T. (1992) *Biochemistry* **31**, 3947-3955.

23. Murphy, K.P. and Freire, E. (1992) *Adv. Prot. Chem*. **43**, 313-361.

24. Makhatadze, G.I. and Privalov, P.L. (1995) *Adv. Prot. Chem*. **47,** 307-425.

25. Myers, J. K., Pace, C. N. and Scholtz, J. M. (1995) *Protein Sci*. **4**, 2138-2148.

26. Spolar, R. S. and Record Jr., M.T. (1994) *Science*, **263**, 777-784.

27. Jin, L., Yang, J. and Carey, J. (1993) *Biochem*. **32**, 7302-7309.

28. Lundback, T. Cairns, C., Gustafsson, J.A., Carlstedt-Duke, J. and Hard, T. (1993) *Biochemistry* **32**, 5074-5082.

29. Merabet, E. and Ackers, G. K., (1995) *Biochemistry* **34**, 8554-8563.

30. Petri, V., Hsieh, M. and Brenowitz, M. (1995) *Biochemistry* **34**, 9977-9984.

31. Morton, C.J. and Ladbury, J.E. (1996) *Prot. Science* **5**, 2115-2118.

32. Berger, C., Jelesarov, I. and Bosshard, H.R. (1996) *Biochemistry* **35**, 14984-14991.

33. Nadassy, K., Wodak, S.J. and Janin, J. (1999) *Biochemistry* **38**, 1999-2017.

34. Kozlov, A.G. and Lohman, T.M. (1999) *Biochem*istry **38**, 7388-7397.

35. Kozlov, A.G. and Lohman, T.M. (2000) *Proteins* Suppl. **4**, 8-22.

36. Beese, L.S., Derbyshire, V. and Steitz, T.A. (1993) *Science*, **260**, 352-355.

37. Eom, S.H., Wang, J. and Steitz, T.A. (1996) *Nature* **382**, 278-281.

38. Li, Y., Korolev, S. and Waksman, G. (1998) *EMBO J.* **17**, 7514-7525.

39. Sterner, R. and Liebl, W. (2001) *Crit. Reviews in Biochem. And Mol. Biol.* **36**, 39-106.

40. Jaenicke, R. (1991) *Eur. J. Biochem*. **202**, 715-728.

41. Jaenicke, R. and Bohm, G. (1998) *Curr. Opin. Struct. Biol.* **8**, 738-748.

42. Hernandez, G., Jenney, F.E., Jr., Adams, M.W.W. and LeMaster, D.M. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 3166-3170.

43. Hosfield, D.J., Frank, G., Weng, Y., Tainer, J.A. and Shen, B. (1998) *J.Biol. Chem*. **273**, 27154-27161.

CHAPTER 5

KLENOW POLYMERASE BINDING TO PRIMED-TEMPLATE DNA EXHIBITS SITE-SPECIFIC THERMODYNAMICS

Introduction

Escherichia coli DNA polymerase 1 (Pol I) was the first DNA polymerase discovered and isolated (1,2), and remains the central model system for studies aimed at understanding DNA replication in general. Functional and mutational studies have shown that the "fingers" region of Klenow bind the single-stranded portion of primed template DNA, while the "thumb" binds the duplex part of primed-template DNA (3-8). Such functional sub-domain mapping of the structural topology of Klenow is largely consistent with the structural data on other Pol I type DNA polymerases for which polymerization site bound DNA co-crystals exist (9,10-14).

In this study we have characterized the temperature dependence of DNA binding by Klenow polymerase in order to understand the thermodynamic driving forces responsible for DNA binding by this polymerase. We find that when compared to the thermodynamics of DNA binding of other DNA binding proteins, Klenow behaves more like a DNA sequence specific binding protein than a non-sequence specific binding protein. Klenow DNA binding is associated with a significant negative ΔCp (heat capacity change), which is manifested as a non-linear temperature dependence of the free energy of binding, and as strong enthalpy-entropy compensation over the temperature range of 5-37^{o}C. Negative ΔCp has been called a "signature" of DNA sequence specific binding (15), and has most often been correlated with the change in the accessible surface area of the protein and the DNA upon complex formation (16,17). Sequence-specific

DNA binding events often bury a relatively large surface area upon complex formation, due to the structurally complementary nature of the interface, and thus are often associated with large negative ΔCps. DNA polymerases are generally considered essentially non-sequence specific DNA binding proteins, although they do exhibit a limited amount of DNA sequence dependent variability of binding affinity (18-20). However, the span between the tightest and weakest binding sequences for a polymerase or other "non-specific" DNA binding proteins (e.g. *E. coli* SSB (21)) are in general on the order of +/- one order of magnitude: a range that is 10^3 to 10^6 fold smaller than the span between sequence-specific and non-specific binding for a transcriptional regulator or restriction endonuclease (15).

Recent studies have shown that the binding affinity of Klenow fragment to a primed-template DNA is significantly reduced by reducing the number of bases in the single-stranded template overhang (6). Similar results are also obtained for DNA binding by T4 DNA polymerase (22). We assayed for sequence specificity of this region of the DNA for binding by changing the bases in the single-stranded template overhang. Our results show that the binding affinities for the DNA sequences tested are very similar with the exception of a poly-A single stranded overhang, which shows slightly weaker affinity as compared to the other sequences. Thus it seems that the binding of the polymerases to DNA is dominated more by the structure of the DNA than by the sequence. Here we propose that analogous to sequence-specific DNA binding proteins, structure-specific DNA binding of Klenow results in a structurally complementary interface between the protein and the DNA resulting in the negative ΔCp observed in this study. Even though a crystal structure of the Klenow-DNA complex in the polymerase

mode is not available, the DNA bound complex for Klentaq DNA polymerase (the large fragment of Taq) shows that the polymerase almost wraps around the DNA encompassing a significant protein-DNA interface.

A temperature dependence study of Taq/Klentaq polymerase DNA binding yielded a similar thermodynamic profile (Chapter 4 and Ref. 23). This study of Klenow polymerase shows that these thermodynamic profiles are not specific to Taq polymerase, but are likely a more universal property of polymerase DNA binding in general.

Experimental Procedures

Materials

All binding studies were performed with the D424A mutant of Klenow known as Klenow exo minus (KF exo-) (See Chapter 2). Table 5.1 shows the different primer-template constructs used. The majority of experiments, including the temperature dependence of binding and the solution structural studies, were carried out with the mixed sequence "13/20mer". The 13/20mer primer template set used is the same as that used for kinetic studies of Klenow DNA binding by Benkovic and associates (18). The longer primer-template pair (63/70mer), that was designed for use at higher temperatures for Taq/Klentaq, was also used this study. The 63/70mer contains the 13/20mer sequence at its primer-template junction, with added "random" sequence selected from the same region of the M13 bacteriophage sequence. DNA oligonucleotides were purchased from IDT (Integrated DNA Technologies Inc.). Fluorescently labeled DNA was labeled at the 5'end of the primer strand with rhodamine-X (ROX) and was purchased directly from Integrated DNA Technology Inc.

Table 5.1. Primer-template DNA Constructs Used for the Experiments

13/20mer "mixed sequence":	5′TCGCAGCCGTCCA3′ 3′AGCGTCGGCAGGTTCCCAAA5′
13/20polyA:	5′TCGCAGCCGTCCA3′ 3′AGCGTCGGCAGGTAAAAAAA5′
13/20 polyT:	5′TCGCAGCCGTCCA3′ 3′AGCGTCGGCAGGTTTTTTTT5′
13/20 polyC:	5′TCGCAGCCGTCCA3′ 3′AGCGTCGGCAGGTCCCCCCC5′
63/70mer:	5′TACGCAGCGTACATGCTCGTGACTGGGATAACCGTGCCGTTTGCCGACTTTCGCAGCCGTCCA3′ 3′ATGCGTCGCATGTACGAGCACTGACCCTATTGGCACGGCAAACGGCTGAAAGCGTCGGCAGGTTCCCAAA5′

Equilibrium DNA Binding Using Fluorescence Anisotropy

Direct equilibrium titrations were performed over a temperature range of 5-37^{O}C using a fluorescence anisotropy binding assay, as described previously in Chapter 3 (24). The DNA used was a 13/20mer primer-template DNA labeled at the 5'end of the primer with rhodamine-X (ROX). For each titration, increasing amounts of Klenow polymerase were titrated into 1nM labeled DNA until saturation. For titrations at temperatures lower than room temperature, nitrogen was flowed through the fluorometer sample chamber at in order to prevent condensation on the cuvette. It was shown previously that the polymerase binds this DNA with 1:1 stoichiometry (Chapter 3 and Ref. 24). The binding isotherms are analyzed with a single-site binding equation as described previously (Chapter 3) to obtain the K_d (dissociation constant) for binding. Unless noted otherwise, all titrations were performed in a buffer containing 10mM Tris, 300mM KCl, 5mM $MgCl_2$, at pH 7.9. Since the pKa of Tris shows a strong temperature dependence, the pH

of the buffers were adjusted at the respective temperatures for all the experiments. pH is adjusted by mixing two buffers containing identical salts and made with Tris-HCl versus Tris base.

The temperature dependence of the free energy of DNA binding was analyzed using an integrated Gibbs-Helmholtz equation with a temperature independent ΔCp as described previously (23).

$$\Delta G_{(T)} = \Delta H_{refT} - T\Delta S_{refT} + \Delta Cp\,[T - T_{refT} - T\ln(T/T_{refT})] \quad \textbf{Eq. 5.1}$$

where $\Delta G_{(T)}$ is the free energy at each temperature, T is the temperature in Kelvin, ΔCp is the heat capacity, and ΔH_{refT} and ΔS_{refT} are the fitted van't Hoff enthalpy and entropy values at any chosen "reference temperature" T_{refT}.

The dependence of equilibrium binding on the DNA sequence within the single stranded region of the primer-template was examined using the 13/20polyA, 13/20polyT and 13/20polyC primer-template constructs shown in Table 5.1, using the fluorescence anisotropy assay. Single-stranded poly-G is known to form stable secondary structures and hence was not examined. For all these titrations, the 5'end of the primer strand was labeled with ROX. All DNA sequence dependence titrations were performed at 25^{O}C in 10mM Tris, 300mM KCl, 5mM MgCl2, pH 7.9.

Isothermal Titration Calorimetry (ITC)

Calorimetric titrations were performed over a temperature range of 10-37^{O}C with unlabeled 13/20mer and 63/70mer DNA in a MicroCal VP-ITC calorimeter. The buffer was identical to the one used for fluorescence anisotropy binding titrations. DNA concentration for the titrations ranged from 1.0 to 2.5uM. The protein concentrations used for the titrations were 30-75uM, such that the starting [DNA]:[Protein] was

maintained at 1:30 for all the titrations. Each titration consisted of an initial 2ul injection (not used for data analysis) followed by 31 subsequent 7ul injections. The heat of dilution of the protein was obtained by titrating protein into the buffer. The actual heat of the reaction was determined after subtracting the heat of dilution of the protein. ITC binding curves were analyzed using the single-site binding equation in the MicroCal Origin software package.

Calorimetric titrations were also performed with 63/70mer DNA at 30^{o}C in buffers with ionization enthalpies (ΔH_{ion}) different from Tris. The DNA concentrations ranged from 2.1-2.5μM. Buffers used were 10mM MOPS and 10mM HEPES at identical salt concentrations and pH values as the Tris based buffer. The enthalpies of binding in these buffers as function of ΔH_{ion} of the buffer (25) were analyzed using the following linear relation in order to determine the effect of buffer ionization on the binding enthalpy:

$$\Delta H_{cal} = n\Delta H_{ion} + \Delta H_{bind} \qquad \textbf{Eq. 5.2}$$

where, ΔH_{cal} is the calorimetric binding enthalpy, n is the number of protons taken up or released by the buffer upon binding, ΔH_{ion} is the ionization enthalpy of the buffer and ΔH_{bind} is the binding enthalpy in the absence of any buffer ionization (26).

Circular Dichroism (CD) Measurements

CD spectra were measured at room temperature (22^{o}C) in an AVIV Model 202 CD spectrophotometer. A dual compartment mixing cuvette was used to record the spectra of protein + DNA before and after mixing. One compartment was filled with 3μM DNA, the other with 3μM Klenow DNA polymerase.

Small Angle X-ray Scattering (SAXS)

SAXS experiments were conducted on synchrotron beam line 1-4 at the Stanford Synchrotron Radiation Research Laboratory (SSRL). Data were collected at a wavelength of 1.488 Å and a sample to detector distance of 0.38m (which is optimal for this beamline). The beam flux was 2 X 10^{10} photons/sec. The sample cell was aluminum and consisted of a 200 μl coin-shaped sample chamber. The path length through the sample was 1.4 mm and the sample chamber was bounded on both sides by circular Kapton windows. SAXS measurements were conducted in the same buffer as used for the binding measurements. Scattering was monitored in 15 minute exposure times for each measurements. Repeated measurements of the same sample yielded the same results, indicating no significant radiation damage occurred during the experiments. The polymerase concentration was 5mg/ml and equimolar concentration of 13/20mer DNA was added for the complex. Previous studies have shown that the determined Rg values do not show any protein concentration dependence within error (27). The scattering data was collected, normalized for dark counts and scattering intensity, and the buffer background scattering was subtracted using the resident analysis programs at the beamline.

SAXS data were analyzed in two ways: 1) Using Guinier plots (28,29) where Rg values were determined from the linear portions of the plots. 2) Using the program GNOM, where an indirect Fourier transform of the experimental scattering curve is implemented to derive the P(r) distance distribution function. The Rg is then calculated from the P(r) function (30). Program default values were used for all input parameters except for Dmax which is altered to obtain the best fit (30).

Results

Temperature Dependence of Equilibrium DNA Binding

Equilibrium binding of Klenow to a 13/20mer primed-template DNA construct was measured over the temperature range of 5-37^{o}C using a fluorescence anisotropy assay. Klenow denatures at ~ 45^{o}C at this pH (31), so 37^{o}C was the highest temperature examined. Figure 5.1 shows representative fluorescence anisotropy titration curves at three temperatures. With increase in temperature the binding affinity ($1/K_d$) increases up until ~25^{o}C, while above this temperature binding gets weaker.

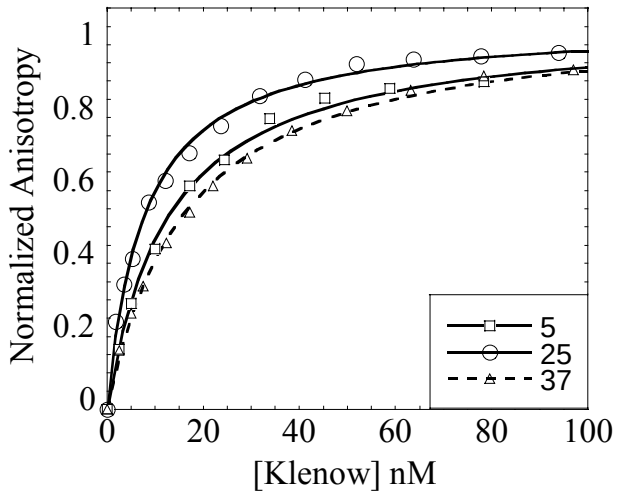


Figure 5.1. Direct equilibrium binding titrations of Klenow DNA polymerase to ROX labeled 13/20mer DNA performed at 5^{o}C (□), 25^{o}C (○), and 37^{o}C (Δ). The binding affinity first increases with increasing temperature up until ~25^{o}C, and then decreases with further increase in temperature.

The temperature dependence of the free energy of binding (ΔG) (Figure 5.2) clearly shows a non-linear behavior with the most favorable free energy of binding at ~30^{o}C. This non-linearity of the temperature dependence of binding is even more

obvious as a van't Hoff plot (Figure 5.2B). This non-linear behavior of the temperature dependence of binding clearly indicates that the binding enthalpy is temperature dependent and that there is an associated ΔCp (heat capacity change) upon binding of Klenow to DNA. If Klenow bound DNA without a heat capacity change (i.e. ΔCp = 0), then the temperature dependence of binding as depicted on both the Gibbs-Helmholtz and van't Hoff plots would be perfectly linear.

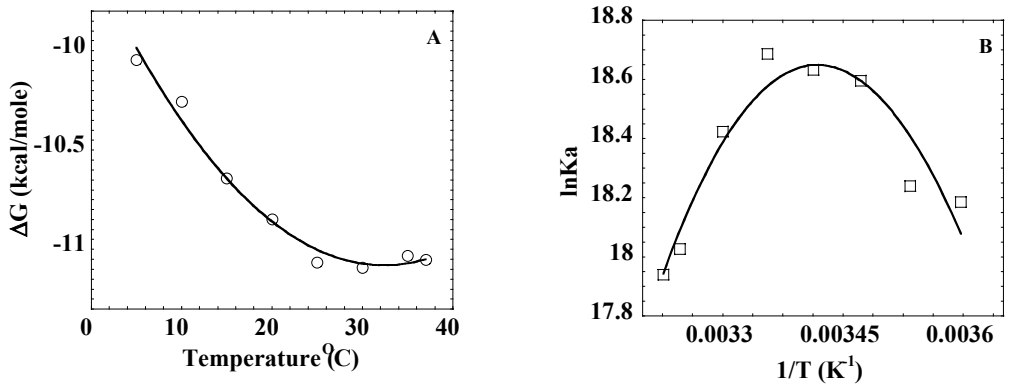


Figure 5.2. A) The temperature dependence of the free energy (ΔG) of the binding of Klenow polymerase to 13/20mer DNA. The line is a fit to the Gibbs-Helmholtz equation, which yields the enthalpy (ΔH_{vH}) and entropy (ΔS_{vH}) at different temperatures for the binding reaction, and heat capacity change (ΔCp). B) The temperature dependence of the K_d of binding of Klenow polymerase to 13/20mer DNA plotted as a van't Hoff plot, and fit to the integrated form of the van't Hoff equation. The data are the same as those in panel A.

To obtain quantitative values for the energetic parameters involved in the DNA binding of Klenow, the temperature dependence of the free energy of binding was analyzed with an integrated Gibbs-Helmholtz equation, as depicted by the fitted line in Figure 5.2. The values of ΔH_{vH}, $T\Delta S_{vH}$, and ΔCp_{vH} are summarized in Table 5.2 and illustrated in Figure 5.3. It is clear that both ΔH_{vH} and $T\Delta S_{vH}$ are strongly temperature dependent. The parallel large magnitude changes in the enthalpy and entropy with temperature, commonly denoted "enthalpy-entropy compensation," results in a

comparatively minimal change of ΔG with temperature. The thermodynamic profile for Klenow binding shown in Figure 5.3 is a profile generally considered characteristic of site-specific protein-DNA binding (15). The ΔCp_{vH} and ΔCp_{cal} for DNA binding of Klenow to 13/20mer DNA (listed in Table 5.2 and Table 5.3) are almost identical, which is an important cross-check of the thermodynamic analysis (32, 33).

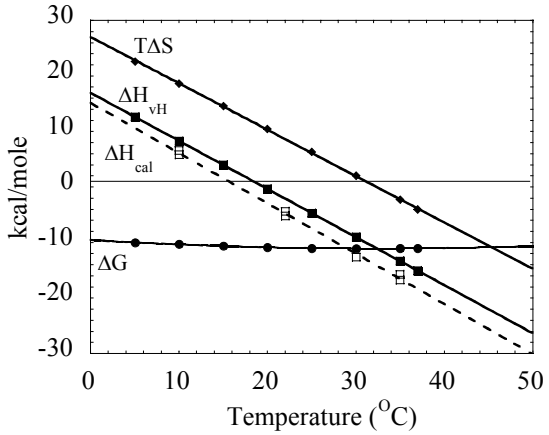


Figure 5.3. Plots showing the changes in the thermodynamic parameters (ΔG, ΔH_{vH}, ΔH_{cal} and $T\Delta S_{vH}$) for binding of Klenow polymerase to 13/20mer DNA as a function of temperature. The figure shows the ΔG (●) values (same data as in Figure 5.2), $T\Delta S_{vH}$ values (♦), ΔH_{vH} values (■), and calorimetric ΔH values (□) determined independently. The binding of Klenow to DNA shows distinct entropy-enthalpy compensation. The binding is enthalpy driven, and entropically opposed at the physiological temperature of ~37°C.

Temperature Dependence of Klenow Binding Monitored Using Isothermal Titration Calorimetry

The enthalpy of DNA binding for Klenow to 13/20mer and to a longer 63/70mer DNA were also measured directly using isothermal titration calorimetry. Titrations were performed from 10-37°C under stoichiometric binding conditions, where Klenow is titrated into DNA. Figure 5.4 shows representative titrations at low and high temperatures along with the fit to a single-site binding equation. The enthalpy of binding

(ΔH_{cal}) obtained from these fits are then plotted as a function of temperature to obtain the calorimetric heat capacity change for binding (ΔCp_{cal}).

Table 5.2. Thermodynamic Parameters of DNA Binding by Klenow DNA Polymerase[a]

Temp (^oC)	K_d (nM)	ΔG (kcal/mol)	ΔH (kcal/mol)	ΔS (kcal/mol K)	$T\Delta S$ (kcal/mol)	ΔCp (kcal/mol K)
5	12.6 ± 0.5	-10.05	+12.63	+0.081	+22.62	
10	14.1 ± 0.7	-10.26	+8.30	+0.066	+18.65	
15	8.4 ± 0.3	-10.64	+3.97	+0.051	+14.62	-0.87±0.13
20	8.1 ± 0.3	-10.85	-0.36	+0.036	+10.5	
25	7.7 ± 0.3	-11.07	-4.69	+0.021	+6.31	
30	10.0 ± 0.3	-11.09	-9.02	+0.007	+2.05	
35	14.8± 0.4	-11.03	-13.25	-0.007	-2.28	
37	16.2± 0.4	-11.05	-15.08	-0.013	-4.03	

[a]K_d and ΔG values are the experimental values determined from titrations at each temperature. The other thermodynamic parameters are calculated from the fit to the Gibbs-Helmholtz equation

Figure 5.5B shows the temperature dependence of ΔH_{cal} for binding of Klenow to 63/70mer DNA. A slightly more negative ΔCp_{cal} is observed for binding to 63/70mer DNA as compared to 13/20mer. The thermodynamic parameters (ΔH_{cal} and ΔCp_{cal}) obtained calorimetrically are summarized in Table 5.3. The temperatures T_H (where ΔH is 0) and T_S (where ΔS is 0) are ~20^{o}C and 32^{o}C respectively for 13/20mer DNA binding by Klenow. The T_H represents the temperature at which K_d is minimum and T_S represents the temperature at which ΔG is minimum.

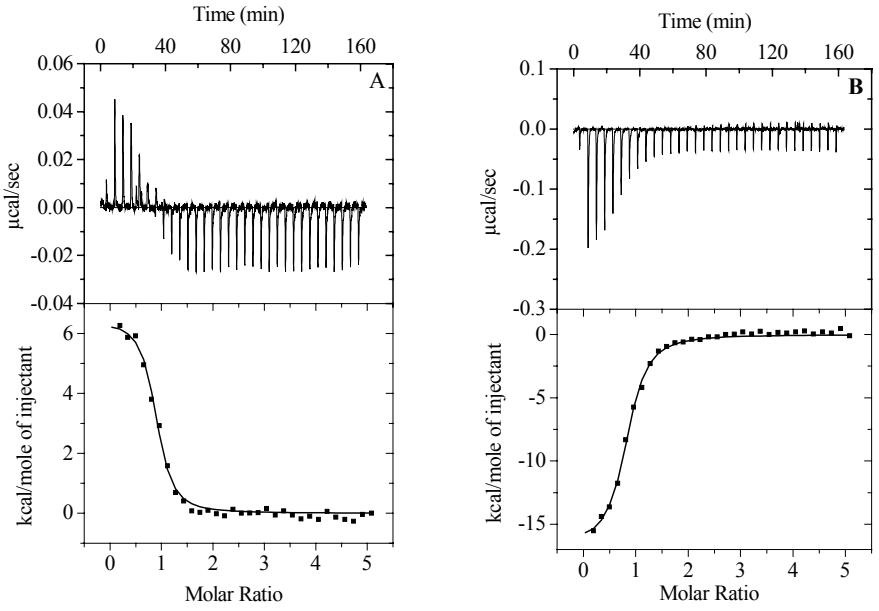


Figure 5.4. Isothermal titration calorimetry for binding of Klenow to unlabeled 13/20mer DNA at 10 and 35°C. The calorimetric enthalpies (ΔH_{cal}) confirm the shift in the thermodynamic driving forces (enthalpy versus entropy) with temperature, see also Figure 5.3.

Table 5.3. Calorimetric ΔH and ΔCp Values for Klenow Binding to 13/20mer and 63/70mer DNA[a]

		13/20mer			**63/70mer**
Temperature(°C)	ΔH_{cal} (kcal/mole)	ΔCp_{cal} (kcal/mol K)	Temperature(°C)	ΔH_{cal} (kcal/mole)	ΔCp_{cal} (kcal/mol K)
10	5.8200		10	+7.85	
10	6.6000		20	-3.99	
22	-4.4120	-0.91±0.02	30	-14.13	-1.22 ± 0.08
22	-5.2500		35	-21.38	
30	-12.660		37	-26.72	
35	-15.730				
35	-16.740				

[a] The ΔH_{cal} values in this table were determined in the presence of 5mM $MgCl_2$.

Assaying for Protonation/Deprotonation Coupled to DNA Binding by Klenow

We also assayed for linked proton effects associated with DNA binding by Klenow. The potential existence of linked proton effects can be evaluated by performing the calorimetric titrations at a particular temperature in buffers with different ionization enthalpies (26). DNA binding titrations were thus performed in Tris (ΔH_{ion}= +11.39kcal/mole), HEPES (ΔH_{ion} = +5.01kcal/mole) and MOPS (ΔH_{ion} = +5.3kcal/mole) buffers (25). Titrations were performed at 30^{o}C with 63/70mer DNA at the same salt concentrations for all three buffers. If protons are taken up or released upon formation of the Klenow-DNA complex, calorimetric measurments should reflect a corresponding heat effect due to the linked stoichiometric protonation or deprotonation of the buffer. Figure 5.5A shows the dependence of ΔH_{cal} as a function of ΔH_{ion} of the buffer. The slope of this plot gives the number of protons released or taken up upon complex formation and the Y-intercept gives the enthalpy of binding in absence of any buffer ionization (26). So, while the directly measured $\Delta H_{binding}$ in Tris buffer at this temperature is –14.1 kcal/mole, the extrapolation of $\Delta H_{binding}$ to zero ΔH_{ion} estimates a $\Delta H_{binding}$ of -17.4 kcal/mole at this temperature, indicating that 3.3 kcal/mole of the measured enthalpy at 30^{o}C is due to a linked protonation effect. The slope of the plot in Figure 5.5A is +0.3, indicating that there is a small net uptake of protons upon binding (25,26). The magnitude of this measured linked proton uptake will be pH dependent (26), so this linkage indicates that at least one ionizable group on the polymerase becomes protonated upon complex formation.

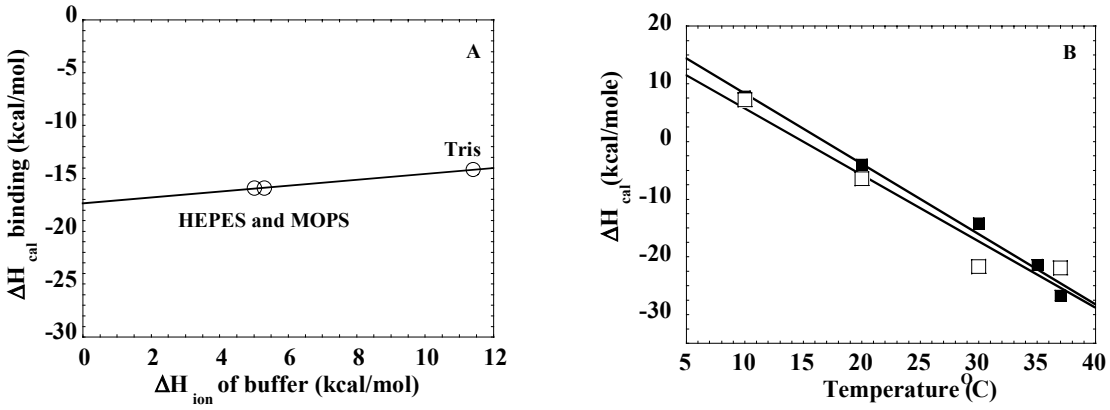


Figure 5.5. A) Calorimetric ΔH values for Klenow-DNA binding at 30°C in buffers with different ionization enthalpies. Measurements were made in HEPES (ΔH_{ion} = 5.01 kcal/mole), MOPS (ΔH_{ion} = 5.3 kcal/mole), and Tris (ΔH_{ion} = 11.39 kcal/mole) (Ref. 25) under identical salt concentrations and pH. The slope of the plot estimates a proton linkage of +0.3 protons (taken up) upon DNA binding (Ref. 25,26). B) Temperature dependences of the calorimetric ΔH of Klenow binding to 63/70mer DNA binding in the presence (■) and absence (□) of $MgCl_2$. The ΔCp values, obtained from the slopes, are nearly identical (ΔCp with Mg^{+2} = -1.22 kcal/mole K; ΔCp without Mg^{+2} = -1.15 kcal/mole K), indicating the ΔCp of binding is not magnesium dependent.

CD and SAXS Measurements of Conformational Changes in the DNA and Protein in Solution

In order to examine the conformational changes of the DNA and the protein upon association in solution, we measured the both CD spectra and the small angle X-ray scattering properties of the free protein and DNA, and the protein-DNA complex. The portions of the CD spectrum corresponding to signal from the protein (between 200-220nm) and from the DNA (~260-290nm) are shown in Figure 5.6 and indicate that there are minor secondary structure rearrangements upon complex formation. CD spectra obtained at room temperature (22°C) and at higher temperatures (37°C) were very similar. The use of a dual compartment mixing cell helps insure that the minor spectral changes observed are real.

Small angle X-ray scattering (SAXS) experiments were also performed, at both 23°C and 37°C, in order to obtain a quantitative measure for the conformational changes of the protein upon DNA binding. A conformational change or linked folding of Klenow

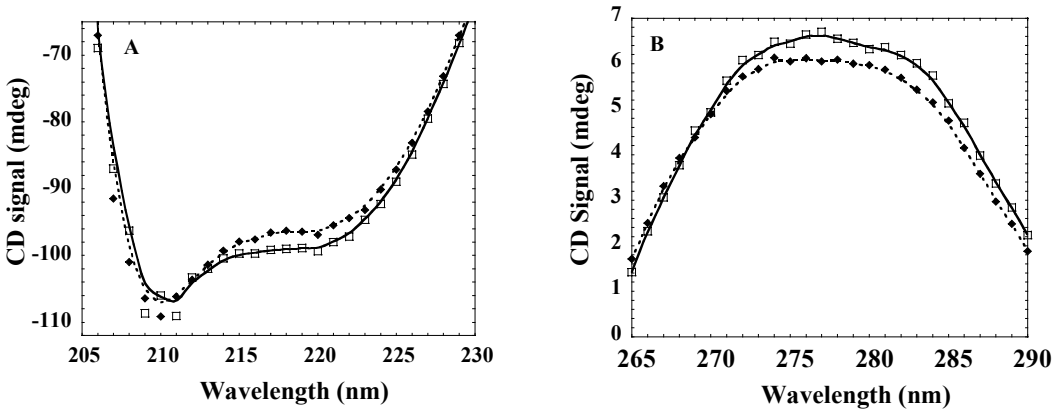


Figure 5.6. Circular dichroism spectra for the Klenow-DNA complex (■) versus the free Klenow + free 63/70mer DNA (□) at 22°C. Spectra were obtained using a dual compartment mixing cuvette and recorded before and after mixing. Shown here are the regions of the spectra showing differences. A) the spectral region from 206 to 228nm corresponding to signal primarily from the protein. B) the spectral region from 265-290nm corresponding to signal from the DNA. Protein + DNA were placed in separate sides of a two compartment mixing cuvette, and spectra were obtained before and after mixing. Spectra were collected in 1nm steps and the lines are simply a smoothed interpolation of the stepwise spectral data. Both the protein and the DNA show slight conformational rearrangements upon complex formation.

upon DNA binding should be manifested by a change in the radius of gyration (R_g) of the protein upon binding. Because the scattering intensity in X-ray scattering is extremely dependent on the size of the scattering particle (intensity is proportional to Rg^6 in Guinier analysis (28,29), the 13/20mer DNA is effectively invisible relative to the protein in these experiments. The Rg values for apo and DNA bound Klenow were determined using both linear (28,29) and non-linear (30) analyses of the SAXS data. Guinier plots at 23°C and 37°C are shown in Figure 5.7, and Rg values are shown in Table 5.4. The increased

intensity (upward shift of the Guinier lines) upon complex formation at both temperatures is likely due to the increased electron density of the protein-DNA complex relative to the free protein. Rg values are obtained from the slopes of these plots. The data at both temperatures indicate that Klenow compacts upon DNA binding, consistent with an induced fit model for DNA binding. The relative compaction appears greater at 37^{o}C than at 23^{o}C, but the relative error of the measurments is also much greater at 37^{o}C.

Structural changes in Klenow are often inferred by extrapolation from structural changes observed in the crystal structures of other type I DNA polymerases, especially Taq polymerase (12,14). Conformational rearrangements upon DNA binding are clearly observable in the comparison of the apo and binary complex structures of Taq polymerase (11,12). The apo structure for Klenow polymerase, however, does not contain side chain data, and as yet there is no crystal structure for a binary complex of Klenow in the polymerization mode. Surface area calculations performed on the apo and DNA bound Taq polymerase crystal structures indicate a slight compaction upon DNA binding in that system (data not shown), corroborating the case for an analogous induced fit compaction of Klenow upon DNA binding.

Assaying the Dependence of Binding Affinity on Template DNA Sequence

Equilibrium titrations were performed with 13/20mer primed template DNA at 25^{O}C with sequences varying in the single stranded template overhang (See Table 5.1). The binding affinities are listed in Table 5.5. The binding affinities for the mixed, poly-T and poly-C are very similar while that of poly-A is slightly weaker.

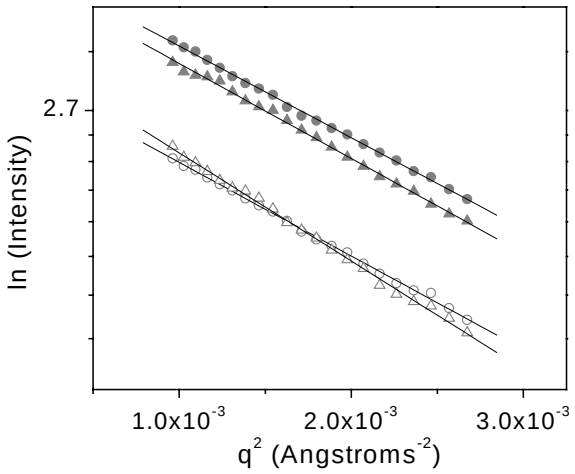


Figure 5.7. Guinier plot of small angle X-ray scattering (SAXS) intensities for free Klenow polymerase (open symbols) and DNA bound Klenow polymerase (closed symbols) at 23°C (circles), and 37°C (triangles) under the same buffer conditions in which the binding experiments were performed. Radius of gyration values for the apo protein and the complex are obtained from the slopes of these lines (slope = $(Rg)^2/3$).

Table 5.4. Radius of Gyration Changes upon DNA Binding to Klenow Polymerase Measured by Small Angle X-ray Scattering

Sample	$Rg_{Guinier}$ (Å)	$\Delta Rg_{Guinier}$[c] (Å)	Rg_{Gnom} (Å)	ΔRg_{Gnom}[c] (Å)
KF apo @ 23°C	28.85 ± 0.13[a]	-0.3 ± 0.18[d]	29.05 ± 0.11[a]	-0.79 ± 0.12[d]
Complex @ 23°C	28.55 ± 0.13[a]		28.26 ± 0.04[a]	
KF apo @ 37°C	31.07 ± 0.29[b]	-1.02 ± 0.93[d]	31.38 ± 0.63[b]	-1.77 ± 1.03[d]
Complex @ 37°C	30.05 ± 0.88[b]		29.61 ± 0.81[b]	

[a] Rg calculated from a single scattering curve ± error of fit.
[b] Mean Rg calculated from 6 scattering curves ± standard deviation.
[c] ΔRg is the $Rg_{Complex}$ minus $Rg_{KF\ apo}$ at each temperature.
[d] Value is ΔRg ± absolute uncertainty.

Table 5.5. The Free Energies of Binding (ΔG) of Klenow to 13/20mer Primed-template DNA with Varying Sequences at the Single-stranded Template Overhang

Sequence of the primed-template DNA	ΔG (kcal/mole)
Mixed sequence	-11.1 ± 0.2
13/20polyT	-11.2 ± 0.2
13/20polyC	-11.1 ± 0.2
13/20polyA	-10.6 ± 0.2

Discussion

Measuring Subtle Changes In Binding Energetics

In this study we have characterized the temperature dependence of DNA binding by Klenow DNA polymerase in order to characterize the thermodynamic driving forces for binding. Our results show that over the temperature range of 5-37^{o}C, the DNA binding affinity of Klenow is distinctly non-linear and has a significant negative ΔCp associated with binding.This results in a temperature dependent enthalpy-entropy compensation of binding. The actual range of K_d values for binding is quite narrow: from 7.7 nM to 16.2 nM. After conversion to free energies at each temperature (ΔG), this corresponds to only about a 1 kcal/mole range from the weakest to the tightest binding. With most DNA binding assays (e.g. filter binding or gel shift) such small changes would be difficult if not impossible to detect, however, as seen from the titration curves in Figure 5.1, such shifts are reliably quantitated using the fluorescence anisotropy assay.

The temperature dependent changes in Klenow's binding affinity are the most subtle (i.e. exibit the narrowest range of change) we have thus far characterized during the course of our solvent and temperature dependent characterizations of different type I

DNA polymerases (23,24). In contrast to the small variation in K_d or ΔG, chemical processes with a significant heat capacity (ΔCp) will exhibit extremely large temperature dependent changes in ΔH and ΔS. For example, for sequence specific DNA binding by BamH1, the range of change of ΔG with temperature is about 1.6 kcal/mole over 2-42^oC temperature range, (15) while the ΔH values for the same reaction change from +25 kcal/mole to –20kcal/mole over the same temperature range (15). The fact that the effective "signal change" from direct measurement of the enthapy is large in magnitude, make the calorimetric measurements of binding enthalpy important corroborating data for ΔG versus temperature measurements. Figure 5.4 shows the calorimetric titration data for Klenow binding at 10 and 35^oC. It is obvious that the enthalpy completely switches sign (from unfavorable to favorable) between the two temperatures. This will only happen if there is a significant ΔCp of binding, and thus this and other ΔH_{cal} data shown in Figures 5.3 and 5.5 corroborate the observed nonlinearity of the ΔG versus temperature measurements in Figure 5.2. Here, the large changes in the value of ΔH_{vH} as a function of temperature is calculated from the small variation in K_d or ΔG with temperature, whereas the variation in ΔH_{cal} with temperature is measured directly. The fact that these two differently derived enthalpies vary nearly identically with temperature yields strong confidence in both the value of ΔCp of binding, and provides confirmation of the anisotropy measurements despite their narrow range. This thermodynamic profile: large difference in the temperature dependences of ΔH and $T\Delta S$ as compared to ΔG, is typical of site-specific DNA binding proteins (34-41,15).

Thermodynamic Driving Forces for DNA Binding by Klenow

The binding of Klenow to primer-template DNA is entropy driven (and enthalpically unfavorable) at lower temperatures and switches to enthalpy driven (and entropically unfavorable) near its physiological temperature of ~37^{o}C. This trend was also observed for DNA binding of Taq DNA polymerase where the binding also switches from entropy driven at low temperatures to enthalpy driven at its physiological temperature of ~75^{o}C (Chapter 4 and Ref. 23). Thus, it seems likely that DNA binding to primed-template DNA is an enthalpy driven process for type I DNA polymerases in general at their respective physiological temperatures.

DNA Binding by DNA Polymerases is Associated with Significant Heat Capacities

A significant negative ΔCp of DNA binding was also found in our recent temperature dependent DNA binding studies of *Thermus aquaticus* DNA polymerase (Taq) (23). This was unexpected since significant values of ΔCp are generally characteristic of site-specific DNA binding by sequence-specific binding proteins (34-41), and DNA polymerase I is generally thought to be mostly non-sequence specific. Since this was the first binding heat capacity data for any DNA polymerase, it was not possible to determine if a significant ΔCp of binding was a unique property of the DNA binding of Taq, or if it was a typical property of DNA polymerases in general. The data of this paper clearly suggest the latter, that DNA polymerases (at least type I polymerases) will likely generally bind primer-template DNA with a significant negative heat capacity. In other words, they will all likely exhibit large excursions of ΔH and $T\Delta S$ with temperature, and switch from entropy to enthalpy driven processes as the

temperature increases, more like sequence-specific DNA binding proteins than non-sequence specific DNA binding proteins.

On the other hand, the non-specific binding complexes of these DNA sequence-specific binding proteins are significantly weaker and generally associated with zero ΔCp (42,43). The magnitude of the ΔCp that we have reported here are more comparable to the values reported for sequence specific DNA binding (15). Although DNA polymerases are not sequence specific binders, they can neither be compared to DNA binding by sequence specific proteins in their non-specific mode. For the polymerases, we believe that it is not the sequence specificity but the recognition of specific DNA structures that contributes to the observed negative ΔCp of binding. Negative ΔCps are also reported for primarily non-specific DNA binders like Sso7d from *Sulfolobus solfataricus* and SSB proteins from *E.coli* (21,44)

Another reason the ΔCp of DNA binding is an important thermodynamic parameter is because it has been shown to be correlated with structural information: specifically the change in accessible surface area upon binding or ΔASA (16,17). Here we report a ΔCp of ~ - 0.9 kcal/mole K for 13/20mer DNA binding by Klenow. The ΔCp for binding to the longer DNA (63/70mer) is somewhat more negative, -1.2 kcal/mole K. The binding interface area calculated from the crystal structure of the editing complex of Klenow bound to DNA (PDB file 1KLN, 45) is ~2722Å^2 (45) (there is no crystal structure of Klenow with DNA in the polymerization mode). The DNA in the crystal structure of the editing complex is a 10/13mer. Almost all nucleotides present in the crystal are in contact with the polymerase (45). At least four different quantitative relationships between ΔCp and the sum of buried non-polar + polar surface area have

been proposed (46-49). A correlation between the ΔCp for binding to the 13/20mer and the reported interface area (45) can be achieved by using the quantitative relationships between ΔCp and ΔASA (46-49). However, all of these equations predict that a relatively high fraction of the buried surface area must be non-polar: the mean of all four relationships predicts that 82% of the buried surface area must be non-polar. Since protein-DNA interfaces are generally much more polar than interiors of proteins, it would seem that other linked processes may also be involved in the generation of the ΔCp of binding of Klenow to DNA.

The relatively high negative ΔCp for the 63/70mer would suggest a larger interface area upon binding to the longer piece of DNA within the context of ΔCp versus ΔASA model. Similar observations were reported in (45) where they showed a positive correlation between interface area and the size of the DNA fragment for DNA binding by DNA polymerases.

DNA Binding by Klenow Does Not Show Any Significant Sequence-specificity

Recent studies have reported the importance of the template overhang for DNA binding of Klenow DNA polymerase (6,8). It has been shown that at least 4 single stranded nucleotides are required in the template overhang for high affinity binding of Klenow and that binding is significantly diminished by reducing the number of nucleotides in the template overhang (6,8). Since our primer-template DNA conforms to the high affinity binding structure we assayed whether the sequence of the template overhang also affected the binding affinity. Our studies show that the binding affinities of Klenow for the varied sequences in the template overhang are generally not significantly different with the possible exception of a poly-A single-stranded region, suggesting that

the sequence in the template overhang does not play a dominant role in the specificity of binding of the primer-template to the polymerase. Poly-A single stranded DNA was also found to differ from other single stranded DNA's in thermodynamics of binding to *E.coli* SSB protein (21). Altered base stacking propensities of poly-A versus other ss-DNA sequences was proposed to be one of the primary reasons for this behavior.

Other Sources of ΔCp for DNA Binding by Klenow

Binding Induced Conformational Changes in the Protein and/or the DNA

We have also looked at the conformational changes in the DNA and/or the protein upon complex formation. The CD spectra of the free and Klenow-DNA complex (Figure 5.6) show a small but observable change in the DNA conformation but no significant change in the secondary structure of the protein. The SAXS experiments shows a small decrease in the Rg of the protein upon DNA binding (see Table 5.4), suggesting a slight compaction of the structure upon complex formation. The interface area calculated by Janin and his coworkers for the co-crystal of Klenow-DNA polymerase estimated a surface area burial of about 2722Å^2, with the ΔASA_{nonpol}=1430Å^2 and ΔASA_{pol}=1292Å^2 (Janin, personal communication). The calculated ΔCp from this estimate is about –0.33kcal/mole K based on ΔCp vs ΔASA correlations reported in Ref 47-50. The binding induced conformational changes observed in SAXS or the CD studies are not significant enough to account for the additional –0.57kcal/mole K of ΔCp observed experimentally.

Restriction of Vibrational Modes

It has been suggested that in addition to the hydrophobic effect, restriction of the vibrational modes of the protein and DNA and the solvent water upon complex formation can contribute to some extent to the observed negative ΔCp (15,51). For EcoR1 and

BamH1, it has been shown that in addition to the hydrophobic effect, the immobilization of water, protein side chains, DNA bases and backbone elements also contribute to the negative ΔCp of DNA binding (15). The co-crystal structure of Klenow with DNA shows more ordered conformation of the protein in certain regions as compared to the apo-structure (52,53). The vibrational state reduction of Klenow upon DNA binding is also suggested by the significant stabilization of the protein upon DNA binding. The free energy of stabilization of the protein will be increased from –4.7 kcal/mole (31) to about -16 kcal/mole (sum of the folding and binding free energies) as a result of DNA binding. This stabilization upon DNA binding is also evidenced by the increase in Tm (melting temperature) of the complex by 10^{o}C as compared to the free protein (the Tm shifts from 45^{o} to 55^{o}C upon binding) (data not shown).

Structure/Function Correlation in the Binding of Klenow to DNA

In a recent review of the still quite sparse collection of thermodynamically characterized non-sequence specific DNA binding proteins, Murphy and Churchill discuss four types or modes of DNA recognition: 1) sequence-specific, 2) non-specific, 3) non-sequence specific, and 4) structure-specific (54). This is, of course, only one possible classification scheme. Here, mode 3, "non-sequence specific" binding is simply the case when a protein has minimal sequence discrimination, but not the large specific versus non-specific site discrimination of sequence-specific proteins. In their review, Murphy and Churchill note that, in general, the sequence-specific binding mode is associated with a large negatvie ΔCp, non-specific binding (mode 2) is associated with near zero ΔCp, "non-sequence" specific binding (mode 3) is associated with a "slightly negative" ΔCp, and that structure-specific binding (mode 4) is as yet not thermodynamically

characterized. Klenow, and our earlier thermodynamic characterization of the ΔCp of DNA binding for the large fragment domain of Taq polymerase, may be the first glimpses of the types of heat capacity changes associated with structure-specific DNA binding. Since structure-specific binding and sequence-specific binding are both variants of "site-specific" binding, it might not be surprising that such binding would be accompanied by a relatively large ΔCp value. Extrapolating from the very few available potential members of this class of DNA binding proteins for which ΔCp information is available (SSB, Sso7d, Taq, and Klenow), it would seem that the magnitude of ΔCp values for DNA structure-specific binding proteins like that of the "non-sequence specific" mode of binding, will partition somewhere between non-specific binders (zero ΔCp) and sequence-specific binding proteins. For SSB (21) and Klenow (18) there clearly exist some DNA sequence preferences in addition to their structure specificity, suggesting that the third and the fourth classes of DNA binding suggested by Murphy and Churchill ("non-sequence-specific" and "structure specific") may involve considerable overlap in thermodynamic character.

More DNA sequence-specific binding proteins have been thermodynamically characterized than non-specific binding proteins. In a recent review of the existing database, Jen-Jacobson and colleagues pointed out that some correlation does exist between higher binding affinity and higher heat capacity (15). It is possible that any DNA binding protein, specific or non-specific, that binds with high enough affinity will exhibit a significant heat capacity change upon binding. Many more systems must be characterized before definitive existence of such a correlation may be discerned. This proposition, however, along with the increasing number of exceptions to the ΔCp-ΔASA

correlation, suggests the possibility of directional correlations between thermodynamic and molecular/structural information. In other words, it is possible that, for example: buried hydrophobic surface area will always exhibit a significant negative ΔCp, but a significant negative ΔCp may not always be quantitatively indicative of buried hydrophobic surface area, and that a lack of a ΔCp may not always be indicative that no hydrophobic surface area is buried. Likewise, DNA sequence specific binding may always exhibit a significant negative ΔCp, but a significant negative ΔCp may not always be indicative of DNA sequence-specific binding, and an absence of a ΔCp may not always be indicative of non-sequence dependent binding. Sturtevant's proposal of several different possible origins for ΔCp (51) has been almost superseded in the scientific literature by a tendency to demonstrate, defend, or attack binary structure-to-ΔCp correlations. Perhaps what the growing database is beginning to show is that some binary correlations, such as ΔCp ≈ ΔASA correlations, and large ΔCp ≈ sequence-specific binding correlations, will empirically dominate protein-DNA interactions, but that a significant subset of protein-DNA interactions will have a more complicated structure to ΔCp relationship. Only continued examination of new and different protein-DNA interactions will elucidate these issues.

References

1. Kornberg, A., Lehman, I.R., Bessman, M.J. and Simms, E.S. (1956) *Biochem. Biophys. Acta*, **21**, 197-198.

2. Lehman, I.R., Bessman, M.J., Simms, E.S. and Kornberg, A., (1958) *J. Biol. Chem*. **233**, 163-170.

3. Catalano, C.E., Allen, D.J. and Benkovic, S.J. (1990) *Biochemistry.* **29**, 3612-3621.

4. Pandy, V.N., Kaushik, N. and Modak, M.J. (1994) *J. Biol. Chem*. **269**, 21828-21834.

5. Minnick, D.T., Astatke, M., Joyce, C.M. and Kunkel, T.A. (1996) *J. Biol. Chem*. **271**, 24954-24961.

6. Turner, R.M., Grindley, N.D.F. and Joyce, C.M. (2003) *Biochemistry.* **42**, 2373-2385.

7. Singh, K. and Modak, M.J. (2003). *J. Biol. Chem*. **278**, 11289-11302.

8. Srivastava, A., Singh, K. and Modak, M.J. (2003) *Biochemistry*. **42**, 3645-3654.

9. Steitz, T.A. (1999) *J. Biol. Chem*. **274,** 17395-17398.

10. Korolev, S., Nayal, M., Barnes, W.M., DiCera, E. and Waksman, G. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 9264-9268.

11. Kim, Y., Eom, S.H., Wang, J., Lee, D.S., Suh, S.W. and Steitz, T.A. (1995) *Nature***. 376**, 612-616.

12. Eom, S.H., Wang, J. and Steitz, T.A. (1996) *Nature* **382**, 278-281.

13. Wang, J., Sattar, K.M.A., Wang, C.C., Karam, J.D., Konigsberg, W.H. and Steitz, T.A. (1997) *Cell,* 89, 1087-1099.

14. Li, Y., Korolev, S. and Waksman, G. (1998) *EMBO J.* **17**, 7514-7525.

15. Jen-Jacobsen, L., Engler, L.E., Ames, J. T., Kurpiewski, M.R. and Grigorescu, A. (2000) *Supramolecular Chemistry* **12**, 143-160.

16. Ha, J.-H., Spolar, R. S. and Record, M.T. (1989) *J. Mol. Biol*. **209**, 801-816.

17. Spolar, R. S. and Record Jr., M.T. (1994) *Science*, **263**, 777-784.

18. Kuchta, R.D., Mizrahi, V., Benkovic, P.A., Johnson, K.A. and Benkovic, S.J. (1987) *Biochemistry* **26**, 8410-8417.

19. Carver, T.E., Jr. and Millar, D.P. (1998) *Biochemistry* **37**, 1898-1904.

20. Thompson, E.H.Z., Bailey, M.F., van der Schans, E.J.C., Joyce, C.M. and Millar, D.P. (2002). *Biochmeistry* **41**, 713-722.

21. Ferrari, M.E. and Lohman, T.M. (1994) *Biochemistry*. 33, 12896-12910.

22. Delagoutte, E. and von Hippel, P.H. (2003) *J. Biol. Chem*. **278**, 25435-25447.

23. Datta, K. and LiCata, V.J. (2003) *Nucleic Acids Research* **31**, 5590-5597.

24. Datta, K. and LiCata, V.J. (2003) *J. Biol. Chem*. **278**, 5694-5701.

25. Lundback, T., van den Berg, S. and Hard, T. (2000) *Biochemistry*, **39**, 8909-8916.

26. Baker, B.M. and Murphy, K.P. (1996) *Biophys. J*. **71**, 2049-2055.

27. Joubert, A.M., Byrd, A.S. and LiCata, V.J. (2003) *J. Biol. Chem*. **278**, 25341-25347.

28. Guinier, A. (1939 *Ann. Phys.* **12**, 166-237.

29. Guinier, A. and Fournet, G. (1955) Small angle scattering of X-rays, Wiley, NY.

30. Semenyuk, A.V. and Svergun, D.I. (1991) *J. Appl. Cryst*. **24**, 537-540.

31. Karantzeni, I., Ruiz, C., Liu, C., and LiCata, V.J. (2003) *Biochem. J*. **374**, 785-792.

32. Liu, Y.F. and Sturtevant, J.M. (1997) *Biophys. Chem*. **64**, 121-126.

33. Horn, J.R., Brandts, J.F. and Murphy, K.P. (2002) *Biochemistry* **41**, 7501-7507.

34. Ladbury, J. E., Wright, J. G., Sturtevant, J. M. and Sigler, P. B. (1994) *J. Mol. Biol.* **238**, 669-681.

35. Merabet, E. and Ackers, G. K., (1995) *Biochemistry* **34**, 8554-8563.

36. Jin, L., Yang, J. and Carey, J. (1993) *Biochemistry* **32**, 7302-7309.

37. Lundback, T. Cairns, C., Gustafsson, J.A., Carlstedt-Duke, J. and Hard, T. (1993) *Biochemistry* **32**, 5074-5082.

38. Petri, V., Hsieh, M. and Brenowitz, M. (1995) *Biochemistry* **34**, 9977-9984.

39. Morton, C.J. and Ladbury, J.E. (1996) *Prot. Science* **5**, 2115-2118.

40. Berger, C., Jelesarov, I. and Bosshard, H.R. (1996 *Biochemistry* **35**, 14984-14991.

41. Lipps, G., Stegert, M. and Krauss, G. (2001) *Nuclec Acids Research*, **29**, 904-913.

42. Takeda, Y., Ross, P.D. Mudd, C.P. (1992) *Proc Natl Acad Sci U S A* **89**, 8180-8184.

43. Heyduk, E., Baichoo, N. and Heyduk, T. (2001) *J. Biol. Chem*. **276**, 44598-44603.

45. Lundback, T., Hansson, H., Knapp, S., Ladenstein, R. and Hard, T. (1998) *J. Mol. Biol*. **276**, 775-786.

46. Nadassy, K., Wodak, S.J. and Janin, J., (1999) *Biochemistry* **38,** 1999-2017.

47. Spolar, R.S., Livingstone, J.R. and Record, M.T. (1992) *Biochemistry* **31**, 3947-3955.

48. Murphy, K.P. and Freire, E. (1992) *Adv. Prot. Chem*. **43**, 313-361.

49. Makhatadze, G.I. and Privalov, P.L. (1995) *Adv. Prot. Chem*. **47**, 307-425.

50. Myers, J. K., Pace, C. N. and Scholtz, J. M. (1995) *Protein Sci*. **4**, 2138-2148.

51. Sturtevant, J.M. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 2236-2240.

52. Beese, L.S., Derbyshire, V. and Steitz, T.A. (1993) *Science* **260**, 352-355.

53. Ollis, D.L, Brick, P., Hamlin, R., Xuong, N.G. and Steitz, T.A. (1985) *Nature* **313**, 762-766.

54. Murphy, F.V. and Churchill, M.E. (2000) *Structure* **8**, R83-R89.

CHAPTER 6

POTENTIAL TEMPERATURE DEPENDENT HEAT CAPACITIES OF BINDING IN PROTEIN-DNA ASSOCIATIONS

Introduction

In Chapter 4 we examined the temperature dependence of DNA binding by Taq DNA polymerase and its 'large fragment' Klentaq, in order to understand the energetics of the reaction. The thermodynamic parameters ΔH, ΔS and ΔCp were derived from Gibbs-Helmholtz analysis of the equilibrium DNA binding as well as measured directly (ΔH_{cal}) by using isothermal titration calorimetry. Our studies showed that DNA binding by Taq/Klentaq DNA polymerases exhibits an enthalpy-entropy compensation due to a negative heat capacity change (ΔCp) of ~ -0.76 kcal/mole K (Chapter 4 and Ref. 1). A negative ΔCp of DNA binding is a common property for the majority of sequence specific DNA binding proteins and is often considered to originate primarily from the burial of non-polar surfaces upon interaction (quantitated by the change in the accessible surface area upon binding, ΔASA) (1-4). In our previous studies (Chapter 4 and 5) we have shown that the negative heat capacity changes of DNA binding by DNA polymerases originate partially, but not exclusively, from the burial of extensive surfaces at the protein-DNA interface and from conformational changes in the protein and the DNA upon binding.

Here we perform a further analysis of the DNA binding thermodynamics of Taq/Klentaq suggesting that the observed heat capacity change of DNA binding exhibits temperature dependence. We compare our results with similar analyses carried out with several other DNA binding proteins (5-7), including one of thermophilic origin (7).

Possible molecular interpretations of the temperature dependent ΔCp of DNA binding are discussed for the DNA binding by Taq/Klentaq DNA polymerases.

Materials and Methods and Data Analysis

The experimental details are reported in Chapter 4. For analysis of the data of Chapter 4 with a temperature dependent ΔCp, the ΔCp was considered to be a linear function of temperature according to the following relation:

$$\Delta Cp_{(T)} = \Delta Cp_{refT} + \Delta\Delta Cp\ (T-T_{refT}) \qquad \textbf{Eq. 6.1}$$

Here $\Delta Cp_{(T)}$ is the heat capacity change at any temperature, T is the temperature in K, ΔCp_{refT} is the heat capacity change at some reference temperature T_{refT} and ΔΔCp is the temperature dependence of ΔCp.

The temperature dependence of the calorimetric enthalpy (ΔH_{cal}) determined from isothermal titration calorimetry (ITC) was thus analyzed with a temperature dependent ΔCp using the following relation (6):

$$\Delta H(_T) = \Delta H_{refT} + \Delta Cp_{refT}(T - T_{refT}) + \Delta\Delta Cp[(\frac{T^2 - T_{refT}{}^2}{2}) - T_{refT}(T - T_{refT})] \qquad \textbf{Eq. 6.2}$$

where $\Delta H_{(T)}$ are the enthalpy changes of binding measured at different temperatures, and ΔCp_{refT} and ΔH_{refT} are the fitted heat capacity change and enthalpy values at any chosen 'reference temperature' T_{refT},

The temperature dependence of the free energy (ΔG) of binding, obtained from the equilibrium binding titrations using fluorescence anisotropy, was analyzed with the Gibbs-Helmholtz equation that also incorporates a ΔΔCp of binding,

$$\Delta G_{(T)} = \Delta H_{refT} + \int_{T_{refT}}^{T} \Delta Cp_T\, dT - T[\Delta S_{refT} + \int_{T_{refT}}^{T} \frac{\Delta Cp_T}{T} dT] \qquad \textbf{Eq. 6.3}$$

$$= \Delta H_{refT} + \Delta Cp_{refT}(T - T_{refT}) + \Delta\Delta Cp[(\frac{T^2 - T_{refT}^{\ 2}}{2}) - T_{refT}(T - T_{refT})] - T\Delta S_{refT}$$
$$- \Delta Cp_{refT} T \ln(\frac{T}{T_{refT}}) - T\Delta\Delta Cp[(T - T_{refT}) - T_{refT} \ln(\frac{T}{T_{refT}})]$$

where, ΔG_T is the free energy change at each temperature, ΔCp_{refT}, ΔH_{refT} and ΔS_{refT} are the fitted van't Hoff heat capacity change, enthalpy and entropy values at any chosen 'reference temperature' T_{refT},

The same sets of data reported in Chapter 4 for equilibrium DNA binding and ITC experiments for binding of Taq/Klentaq DNA polymerase to 63/70mer DNA were used for this analysis. The temperature range selected for this study is between 10-60^{O}C.

Circular Dichroism (CD) Measurements

CD spectra were measured at 10, 22 and 30^{o}C in an AVIV Model 202 CD spectrophotometer. A dual compartment mixing cuvette (Starna Cells), was used to record the spectra of protein + DNA before and after mixing. One compartment was filled with 3 μM DNA, the other with 3μM Klentaq. See Chapter 4 for detailed procedure.

Results

Figure 6.1 shows the temperature dependence of the calorimetric enthalpy (ΔH_{cal}) of DNA binding by Taq and Klentaq. Non-linear analysis of the ΔH_{cal} versus T for Taq using equation 6.2 yields the calorimetric $\Delta\Delta Cp_{cal}$ (listed in Table 6.1). The absolute ΔΔCp value estimated for Taq is somewhat smaller than that for Klentaq.

The temperature dependence of the free energy of DNA binding by Taq/Klentaq is shown in Figure 6.2 where the solid line shows the fit to the Gibbs-Helmholtz equation with a temperature dependent ΔCp (Eq. 6.3). The binding parameters at each temperature,

derived from the Gibbs-Helmholtz analysis, are listed in Table 6.2. Although the absolute ΔΔCp obtained from Gibbs-Helmholtz analysis differs somewhat from the $\Delta\Delta Cp_{cal}$, the general trend is similar, and again the ΔΔCp for Klentaq is more negative than Taq.

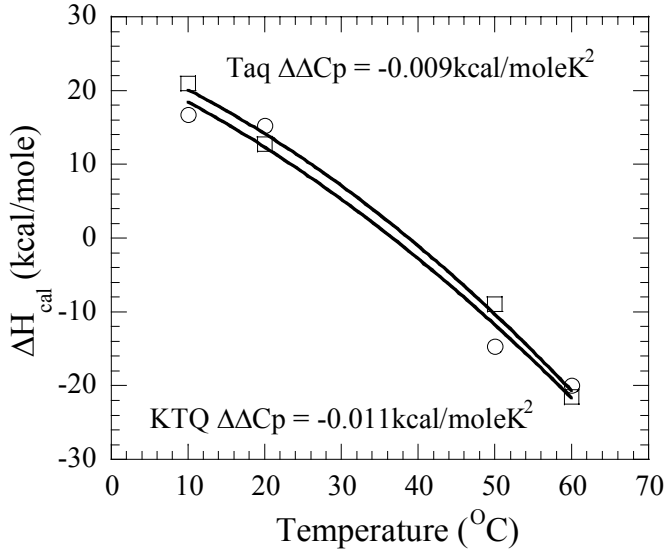


Figure 6.1. Temperature dependencies of the calorimetric ΔH for DNA binding by Taq (○) and Klentaq (□) DNA polymerases, analyzed with a temperature dependent ΔCp. Data are those from Figure 4.5 in Chapter 4. The solid lines are the fits to Eq. 6.2 in the text, that includes a temperature dependent ΔCp.

Table 6.1. Calorimetric ΔH and ΔΔCp Values for Taq and Klentaq

	Taq		**Klentaq**	
Temperature(°C)	ΔH_{cal} (kcal/mole)	$\Delta\Delta Cp_{cal}$ (kcal/mol K)	ΔH_{cal} (kcal/mole)	$\Delta\Delta Cp_{cal}$ (kcal/mol K)
10	+16.73		+20.93	
20	+15.21	-0.009± 0.02	+12.75	-0.011± 0.01
50	-14.64		-8.93	
60	-19.95		-21.55	

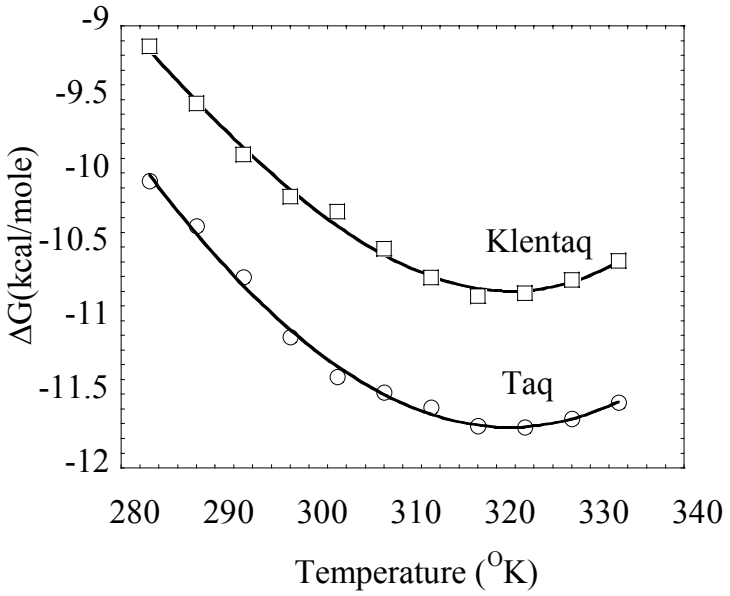


Figure 6.2. Gibbs-Helmholtz plot showing the temperature dependence of the free energy (ΔG) of DNA binding for Taq (○) and Klentaq (□), analyzed with a temperature dependent heat capacity change. Lines are the fits to the expanded Gibbs-Helmholtz equation (Eq. 6.3) as described in Materials and Methods.

Table 6.2. Thermodynamic Parameters of DNA Binding by Taq and Klentaq DNA Polymerases Analyzed with a Temperature Dependent ΔCp

Taq

Temp (°C)	K_d (nM)	ΔG (kcal/mol)	ΔH (kcal/mol)	ΔS (kcal/mol K)	TΔS (kcal/mol)	ΔCp (kcal/mol K)	ΔΔCp (kcal/mol K^2)
10	17.3 ± 0.6	-10.05	+13.87	+0.084	+23.77	-0.513	
15	13.8 ± 0.5	-10.36	+11.21	+0.075	+21.63	-0.553	
20	10.3 ± 0.4	-10.71	+8.35	+0.065	+19.1	-0.592	
25	7.1 ± 0.2	-11.11	+5.29	+0.055	+16.36	-0.632	
30	6.2 ± 0.3	-11.38	+2.03	+0.044	+13.33	-0.672	-0.008±0.009
35	7.1± 0.2	-11.49	-1.44	+0.033	+10.07	-0.712	
40	8.1± 0.3	-11.59	-5.09	+0.021	+6.54	-0.752	

(table cont'd)

45	8.9 ± 0.4	-11.71	-8.95	+0.009	+2.77	-0.792	
50	11.7 ± 0.6	-11.72	-13.01	-0.004	-1.29	-0.832	
55	16.8 ± 0.6	-11.67	-17.27	-0.017	-5.61	-0.872	
60	26.0 ± 1.4	-11.56	-21.73	-0.031	-10.19	-0.911	

Klentaq

Temp (^oC)	K_d (nM)	ΔG (kcal/mol)	ΔH (kcal/mol)	ΔS (kcal/mol K)	TΔS (kcal/mol)	ΔCp (kcal/mol K)	ΔΔCp (kcal/mol K^2)
10	87.7± 1.8	-9.14	+10.70	+0.070	+19.81	-0.184	
15	59.0 ± 0.8	-9.53	+9.54	+0.066	+19.01	-0.281	
20	43.3± 1.1	-9.87	+7.9	+0.060	+17.58	-0.377	
25	35.6 ± 1.1	-10.16	+5.77	+0.053	+15.79	-0.474	
30	39.8 ± 1.4	-10.26	+3.16	+0.045	+13.64	-0.570	-0.019±0.009
35	34.6 ± 1.1	-10.51	+0.07	+0.034	+10.47	-0.666	
40	33.3 ± 1.3	-10.71	-3.50	+0.023	+7.20	-0.763	
45	35.8 ± 1.5	-10.83	-7.56	+0.010	+3.18	-0.859	
50	48.3 ± 1.5	-10.81	-12.1	-0.004	-1.29	-0.955	
55	71.4 ± 3.0	-10.72	-17.11	-0.019	-6.23	-1.05	
60	111.4 ± 5.7	-10.59	-22.61	-0.036	-11.99	-1.15	

Figure 6.3 shows the temperature dependence of the ΔH_{cal} for DNA binding by Taq/Klentaq and other DNA binding proteins for which a temperature dependence of the ΔCp of binding have been reported (5-7). The temperature dependence of the ΔCp of DNA binding by Taq/Klentaq DNA polymerases and the other proteins that exhibit ΔΔCp for DNA binding are shown in Figure 6.4. Like Taq/Klentaq, the ΔCp of DNA binding for all these proteins change in same direction and with similar magnitudes as the temperature increases.

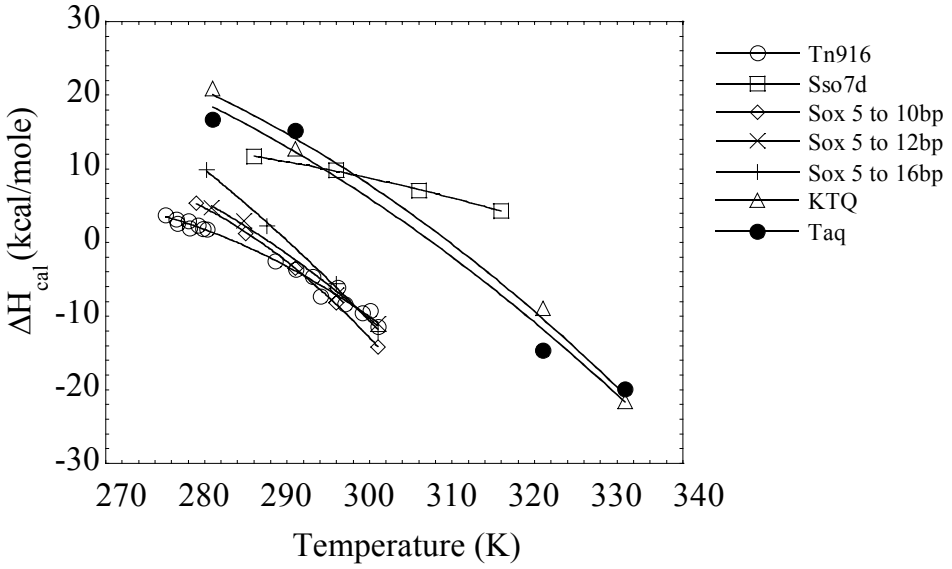


Figure 6.3. Temperature dependence of the calorimetric enthalpy change (ΔH_{cal}) for Taq and Klentaq and other DNA binding proteins exhibiting a temperature dependent ΔCp of DNA binding.

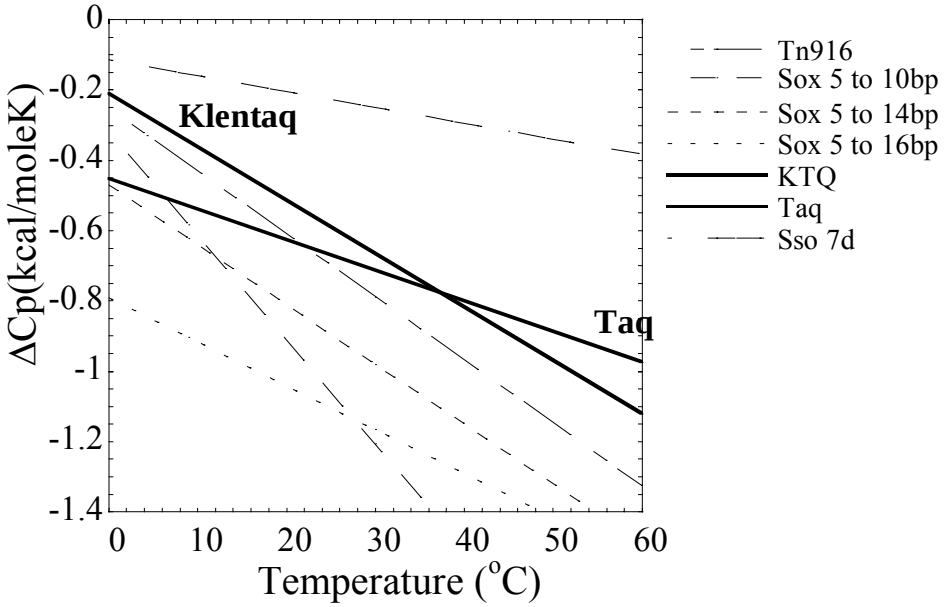


Figure 6.4. Temperature dependence of the ΔCp of DNA binding for Taq and Klentaq DNA polymerases (solid lines) and other DNA binding proteins (dashed lines) (Ref. 2-4). The values are reported in Table 6.3.The $\Delta\Delta Cp$ for Taq and Klentaq are the means of the calorimetric and van't Hoff measurements.

Table 6.3. Calorimetric ΔΔCp Values for Other Protein-DNA Interactions

Protein-DNA interaction	$\Delta\Delta Cp_{cal}$ (kcal/mol K^2)
Sso-7d/ poly(dGdC) [a]	-0.005 ± 0.002
Sox 5 HMG box/ 10bp specific DNA [b]	-0.03 ± 0.004
Sox 5 HMG box/ 12bp specific DNA [b]	-0.017 ± 0.025
Sox 5 HMG box/ 16bp specific DNA [b]	-0.012 ± 0.016
Tn916/specific DNA [c]	-0.018 ± 0.006
Taq/63/70merDNA [d]	-0.0087 ± 0.024
Klentaq / 63/70mer DNA [d]	-0.015 ± 0.011

[a] Ref.7, [b]Ref.5, [c]Ref.6, [d]measured in this study and averages of the calorimetric and van't Hoff measurements.

Figure 6.5 shows the CD difference spectra (complex-free) corresponding to the signal from the protein, for the binding of Klentaq to 63/70mer DNA in buffer conditions similar to that used for the binding reactions.

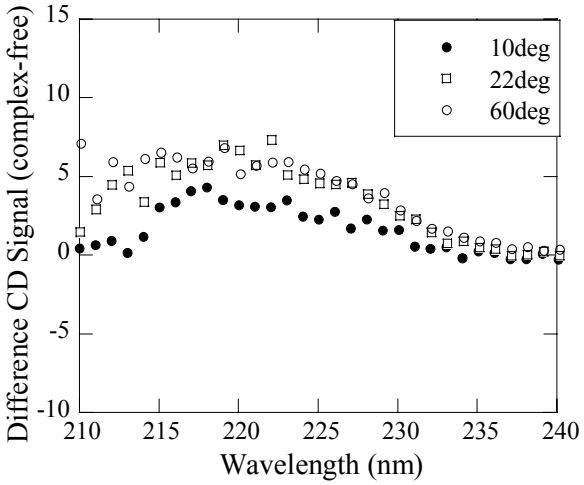


Figure 6.5. The CD difference spectra of Klentaq-DNA complex – free components at 10 (●), 22 (□) and 60°C (○).

Discussion

One of the goals of biothermodynamics is to be able to predict molecular details for any macromolecular process based primarily on energetic information. In the search for the principles that would link thermodynamics to structural information, the heat capacity change (ΔCp) has attracted special attention as it has frequently been found to be correlated with the changes in the solvent accessible surface area (ΔASA) associated with the molecular process (1-4). Protein-DNA interactions are usually associated with negative ΔCp that has been ascribed to the ΔASA associated with the burial of macromolecular surfaces at the protein-DNA interface. Consequently, an accurate understanding of ΔCp is crucial for an accurate structural interpretation of ΔCp.

For DNA-protein interactions, the ΔCp of complex formation ($\Delta Cp_{binding}$) is the difference between the absolute heat capacities (Cp) of the complex and the individual reactants (protein and DNA),

$$\Delta Cp_{binding} = Cp_{complex} - (Cp_{DNA} + Cp_{protein})$$

This would provide a correct estimate of the $\Delta Cp_{binding}$, but would be independent of temperature only if the heat capacities of the components of the reaction do not change over the considered temperature range, or if they change in such a way that the difference between $Cp_{complex}$ and $(Cp_{DNA} + Cp_{protein})$ does not change over the considered temperature range (8). However, if the partial heat capacities of the individual components exhibit different temperature dependence, the $\Delta Cp_{binding}$ will no longer be constant with temperature, but instead would show temperature dependence that will be defined by the temperature dependencies of the individual reacting components,

$$\frac{\partial \Delta Cp_{binding}}{\partial T} = \frac{\partial Cp_{complex}}{\partial T} - (\frac{\partial Cp_{DNA}}{\partial T} + \frac{\partial Cp_{protein}}{\partial T})$$

For protein-DNA interactions, the non-linear temperature dependencies of the calorimetric enthalpy change (ΔH_{cal}) of binding is considered to be an indication of the presence of temperature dependent $\Delta Cp_{binding}$ ($\partial \Delta Cp_{binding}/\partial T$ or $\Delta\Delta Cp$) (5,6).

Figure 6.1 shows the temperature dependencies of the ΔH_{cal} for DNA binding by Taq/Klentaq. Although, the linear analysis of this data, as had been performed previously (see Chapter 4), yields a reasonable estimate of the ΔCp of binding, the slight curvature in the plot suggests the presence of a $\Delta\Delta Cp$ that was not considered in our previous analysis. Non-linear analysis of the ΔH_{cal} versus T plot here, provides a $\Delta\Delta Cp$ associated with DNA binding by Taq/Klentaq (Table 6.1). Analysis of the Gibbs-Helmholtz plot with a $\Delta\Delta Cp$, as shown in Figure 6.2, also yields an associated $\Delta\Delta Cp$ for DNA binding by Taq/Klentaq. The similar magnitudes of the $\Delta\Delta Cp$'s for the two proteins from the two different techniques reinforce their plausibility.

Temperature dependent ΔCp values have been observed previously for sequence specific DNA binding by Sox-5 HMG box protein from mouse (5), Tn916 from *E.coli* (6) and non-sequence specific binding by Sso7d protein from *Sulfolobus* bacteria (7). The respective $\Delta\Delta Cp$ values are listed in Table 6.3. Figure 6.3 and 6.4 illustrate the ΔH_{cal} versus T and the ΔCp versus T behavior for all of these systems. It is notable that the estimated $\Delta\Delta Cp$, observed for DNA binding by Taq/Klentaq, are similar to those observed for the other DNA binding proteins. The similar magnitudes for all these systems, and the fact that ΔCp decreases with temperature in all of these systems, again reinforces their collective plausibility. Furthermore, $\Delta\Delta Cp$ values were obtained in

several different ways in these studies. In our study, the non-linear analysis of the Gibbs-Helmholtz plot, using equation 6.3 and analysis of ΔH_{cal} with Eq.6.2 yield a similar estimate of the ΔΔCp. For Sso7d, the ΔΔCp was also confirmed by Gibbs-Helmholtz analysis (7). For the DNA binding by Sox-5 HMG box protein and Tn916, listed in Table 6.3, parallel determination of the ΔΔCp of binding was performed by measuring the $(\partial \Delta Cp_{DNA}/\partial T)$, $(\partial \Delta Cp_{protein}/\partial T)$ and $(\partial \Delta Cp_{complex}/\partial T)$ using differential scanning calorimetry (DSC) (5,6).

The temperature dependence of the ΔCp reported for protein-DNA interactions has been attributed primarily to the non-parallel " thermal" or "conformational fluctuations" of the reacting components over the temperature range where binding is studied (5,6). These conformational fluctuations are fundamentally different from the binding-induced conformational changes reported for protein-DNA interactions (9). For e.g., in the unbound state, one of the 'minor wings' of the Sox-5 protein is partially unfolded in solution, and refolds to be able to bind the DNA. This partial unfolding of the protein, that is associated with a heat capacity change with temperature $(\partial \Delta Cp_{protein}/\partial T)$, constitutes the "thermal fluctuations" referred above. Similar thermal fluctuations are also observed for Tn916 (6). In these studies, the authors have shown that correcting for these restricted thermal motions of the reacting components would yield a ΔCp of binding that is essentially temperature independent (5,6). Although this correction provided a better correlation between the experimental ΔCp and the ΔCp calculated from x-ray crystal structures (based on ΔCp versus ΔASA correlations), the importance of the observed ΔΔCp of binding has also been emphasized (6). Here the authors have pointed out that considering the presence of ΔΔCp for binding is important since it represents the

association of the protein and the DNA in their real conformational states, including thermal fluctuations, conformational changes and partial unfolding (6).

Although our present data does not allow us to deconvolute the observed ΔΔCp for DNA binding by Taq/Klentaq into exact contributions from the partial heat capacities of the reacting species, the similarity of the ΔΔCp value with other DNA binding proteins, as shown in Table 6.3, suggests that similar "thermal or conformational fluctuations" might also be associated with the DNA binding by these polymerases.

Probable Molecular Origins of the Temperature Dependence of the ΔCp of DNA Binding

In 1977, J.M. Sturtevant identified several potential molecular origins of ΔCp for macromolecular processes (10). Here, I have discussed the probable contributions of each of these processes for the DNA binding by Taq/Klentaq.

Hydrophobic Effects

This results from the formation of cages of structured water of abnormally high heat capacity and low entropy around the nonpolar groups of the macromolecules (10). The burial of the nonpolar surfaces at the interface due to complex formation, results in the dehydration of these nonpolar surfaces, thereby lowering the heat capacity of the complex. This results in a negative ΔCp upon complex formation and can be correlated with the change in the accessible surface area (ΔASA) upon binding (1-4). The reported ΔASA value due to formation of the protein-DNA interface for the DNA binding of Taq is 2530$Å^2$ (11), and this will contribute partially to the negative ΔCp observed in this study.

Electrostatic Charges

It has been shown that the creation of positive and negative pairs of charges in aqueous solution decreases the ΔCp of the reacting solutes (10). In view of the strong tendency of all charged groups to be immersed in the aqueous solvent rather than the relatively low dielectric interior of a macromolecule, it was proposed that the change in the exposure of the charged groups to the solvent does not change much in most macromolecular processes (10), thereby not contributing significantly to the ΔCp.

However, in contrast to the interior of proteins, the protein-DNA interface is relatively more charged and polar due to the presence of the DNA phosphates and the positively charged residues of the protein. The ratio of $\Delta ASA_{pol}/\Delta ASA_{nonpol}$ for protein folding is about 0.59 (9) whereas for DNA binding by Taq, the calculated $\Delta ASA_{pol}/\Delta ASA_{nonpol}$ in the protein-DNA interface is about 0.87 (J. Janin, personal communication). Thus, the electrostatic interactions in the protein-DNA interface are also likely to partially contribute to the observed negative ΔCp.

Hydrogen Bonds

The change in the net extent of hydrogen bonds during a reaction will also be reflected in the ΔCp of the reaction (10). However, it was proposed that in most protein reactions the energy penalty for less than maximal hydrogen bonding would insure that there is no significant change in the extent of hydrogen bonding (10), therefore not contributing significantly to the ΔCp.

Intramolecular Vibrations

The changes in the number of easily excitable internal vibrational modes of macromolecules are also considered to be a possible source of the heat capacity change in

macromolecular processes (10). Restriction of these vibrational degrees of freedom upon protein folding or protein-DNA interaction will contribute to the negative heat capacity changes observed for these interactions (10,12).

The co-crystal structure of Klentaq with DNA shows "disordered to ordered transition" in some regions of the protein (13), as compared to the apo- structure (14), indicating binding induced folding of the protein. This may also likely contribute to the restriction of the vibrational degrees of freedom of the protein upon complex formation, and thus contributing partially to the negative ΔCp of binding.

The molecular processes described above will contribute to the ΔCp's observed for protein-DNA interactions. However, it seems that the extent of each these contributions will vary for different protein-DNA systems. The temperature dependence of any of these processes will also be reflected in the observed ΔCp, resulting in a temperature dependent heat capacity change.

Among the molecular processes discussed above, the hydrophobic effect, reflected in the burial of nonpolar surfaces, is generally considered to be the most dominating contribution to the ΔCp. For DNA binding by Taq, if we consider that the $\Delta ASA_{pol}/\Delta ASA_{nonpol}$ =0.87, then using the ΔCp vs ΔASA correlation in Ref. 1, our observed ΔΔCp predicts that the $\Delta ASA_{nonpolar}$ will change from 2590Å^2 at 10^{o}C to 4601Å^2 at 60^{o}C. Our CD difference spectrum (complex-free), corresponding to the signal from the protein, at temperatures between 10-60^{o}C, shows a small conformational change between 10 and 25^{o}C but no noticeable change above that (See Figure 6.4). The CD spectrum for the DNA does not show any observable change over the entire temperature range of 10-60^{o}C (data not shown). Further, the radii of gyration (Rg values) of the free

and DNA bound polymerases, measured by small angle x-ray scattering (SAXS), do not change significantly over the temperature range of 23-60^{o}C (data not shown). These spectral and structural data suggest that the temperature dependent conformational changes associated with the DNA binding by Taq/Klentaq can account for some of the change in ΔASA, but are not significant enough to account for all the predicted ΔASA at the higher temperatures if the origin of ΔCp were primarily due to hydrophobic effects.

Our preliminary studies show that DNA binding by Klentaq at 25^{o}C is associated with a significant decrease in the volume of water associated with the macromolecular surfaces (See Appendix 1), indicating a net release of water upon binding. However, the estimated water release is much larger number than would be expected on the basis of only hydrophobic effects (i.e. ΔASA). Thus, as has been proposed for protein folding reactions (10), it may be possible that part of this aqueous volume change contributes to the decrease in the internal vibrational modes of the macromolecules upon binding, thereby contributing to the negative ΔCp, but not necessarily to the burial of additional hydrophobic surfaces. Temperature dependence of this phenomenon would contribute to the temperature dependence of ΔCp.

As has been pointed out previously, and also evident from these results, Taq/Klentaq bind DNA with high affinity at temperatures as low as 5-10^{O}C, although they are catalytically almost inactive at room temperature (15,16). The low temperature DNA binding is also observed for other thermophilic proteins (7,7-19), including Sso7d mentioned in this study. It is generally believed that thermophilic proteins are more rigid at room temperatures (20,21), which is correlated with a decreased ability to undergo the conformational fluctuations required for catalytic activity, at low temperatures. It seems

that temperature dependent configurational vibrations of Taq/Klentaq, as proposed here, are also necessary for the catalytic activity at higher temperatures.

References

1. Spolar, R.S., Livingstone, J.R. and Record, M.T. (1992) *Biochemistry* **31**, 3947-3955.

2. Murphy, K.P. and Freire, E. (1992) *Adv. Prot. Chem*. **43**, 313-361.

3. Makhatadze, G.I. and Privalov, P.L. (1995) *Adv. Prot. Chem*. **47**, 307-425.

4. Myers, J. K., Pace, C. N. and Scholtz, J. M. (1995) *Protein Sci*. **4**, 2138-2148.

5. Privalov, P.L., Jelesarov, I., Read, C.M., Dragan, A.I. and Crane-Robinson, C. (1999) *J. Mol. Biol*. **294**, 997-1013.

6. Milev, S., Gorfe, A.A., Karshikioo, A., Clubb, R.T., Bosshard, H.R. and Jelesarov, I. (2003) *Biochemistry*, **42**, 3481-3491.

7. Lundback, T., Hansson, H., Knapp, S., Ladenstein, R. and Hard, T. (1998) *J. Mol. Biol*. **276**, 775-786.

8. Liggins, J.R. and Privalov, P.L. (2000) *Proteins: Str. Func. and Gen. Suppl*. 4, 50-62.

9. Spolar, R. S. and Record Jr., M.T. (1994) *Science*, **263**, 777-784.

10. Sturtevant, J.M. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 2236-2240.

11. Nadassy, K., Wodak, S.J. and Janin, J., (1999) *Biochemistry* **38**, 1999-2017.

12. Jen-Jacobsen, L., Engler, L.E., Ames, J. T., Kurpiewski, M.R. and Grigorescu, A. (2000) *Supramolecular Chemistry* **12**, 143-160.

13. Li, Y., Korolev, S. and Waksman, G. (1998) *EMBO J.* **17,** 7514-7525.

14. Korolev, S., Nayal, M., Barnes, W.M., Di Cera, E. and Waksman, G. (1995) *Proc. Natl. Acad. Sci. U.S.A.* 92, 9264-9268.

15. Lawyer, F.C., Stoffel, S., Saiki, R.K., Chang, S.Y., Landre, P.A., Abramson, R.D. and Gelfand, D.H. (1993) *PCR Methods Appl.* **2**, 275-287.

16. Hogrefe, H.H., Cline, J., Lovejoy, A.E. and Nielson, K.B. (2001) *Methods Enzymol*. **334**, 91-116.

17. McAfee, J. G., Edmondson, S. P., Zegar, I. and Shriver, J. W. (1996) *Biochem*. **35**, 4034-4045.

18. Lipps, G., Stegert, M. and Krauss, G. (2001) *Nuclec Acids Research*, **29**, 904-913.

19. Hosfield, D.J., Frank, G., Weng, Y., Tainer, J.A. and Shen, B. (1998) *J.Biol. Chem*. **273**, 27154-27161.

20. Jaenicke, R. (1991) *Eur. J. Biochem*. **202**, 715-728.

21. Jaenicke, R. and Bohm, G. (1998) *Curr. Opin. Struct. Biol.* **8**, 738-748.

CHAPTER 7

CONCLUDING SUMMARY

In this dissertation I have presented a thermodynamic characterization of DNA binding by the type I DNA polymerases from *Thermus aquaticus* (Taq/Klentaq) and *E.coli* (Klenow fragment). This is the first in-depth characterization of the detailed thermodynamics (i.e. ΔG, ΔH, ΔS and ΔCp) of DNA binding for any DNA polymerase. The salt dependence of DNA binding by Klenow and Taq/Klentaq DNA polymerases has also been studied as a function of [KCl] and [$MgCl_2$].

Our studies show that the binding of the two different species of polymerases (Taq/Klentaq versus Klenow) occurs with sub-micromolar affinities in very different salt concentration ranges. As a consequence, at similar [KCl] the binding of Klenow is about 3kcal/mol (150X) tighter than the binding of Taq/Klentaq to the same DNA. Linkage analysis of the [KCl] dependence reveals a net release of 2-3 ions upon DNA binding of Taq/Klentaq, and a net release 4-5 ions upon DNA binding of Klenow. DNA binding of Taq at higher temperature (60^{o}C) only slightly decreases the linked ion release. Linkage analysis of DNA binding as a function of increasing [$MgCl_2$] indicates that formation of the protein-DNA complex in both species of polymerase is ultimately linked to the release of approximately one Mg^{+2} ion. However, the $MgCl_2$ dependence for Klenow polymerase, but not Klentaq, shows two distinct phases, with a small Mg^{+2} uptake at $MgCl_2$ concentrations below 10mM, and Mg^{+2} release dominating the dependence at $MgCl_2$ concentrations above 10mM. In the presence of 10mM EDTA, both species of polymerase can still bind DNA, but their DNA binding affinity is significantly diminished, Klenow more so than Klentaq. It is not yet determined which, if any, of these

functional differences are correlated with the ability of Taq/Klentaq polymerase to function at such dramatically higher temperatures than Klenow.

DNA binding by these polymerases was also studied as a function of temperature in order to understand the thermodynamic driving forces involved in DNA binding by the polymerases. Direct equilibrium titrations using a fluorescence anisotropy assay were performed across a temperature range of 5-70^{o}C for Taq/Klentaq, and 5-37^{o}C for Klenow. Temperature dependences of the free energy (ΔG) of DNA binding by Taq/Klentaq and Klenow show distinct curvature due to the presence of a negative heat capacity change (ΔCp) associated with binding. For Taq/Klentaq the binding is most favorable at ~ 40-50^{O}C while for Klenow the most favorable binding is observed at ~ 25-30^{O}C. Nonlinear Gibbs-Helmholtz analysis of the ΔG vs T dependence gives the van't Hoff ΔH, ΔS and ΔCp of DNA binding. It is notable that like sequence specific DNA binding proteins, DNA binding by these polymerases show enthalpy- entropy compensation across the entire temperature range studied, with the binding being entropy driven at lower temperatures and switching to enthalpy driven at higher temperatures. For both the polymerases DNA binding is enthalpy driven at their respective physiological temperatures. DNA binding titrations of these polymerases using isothermal titration calorimetry confirmed the thermodynamic parameters quantitated from van't Hoff analysis. An unusual finding was that the DNA binding activity of Taq/Klentaq is maintained at temperature as low as 5^{o}C even though like other thermophilic proteins it loses catalytic activity below room temperature. The ΔCp of DNA binding by Taq/Klentaq is ~ –0.7 to -0.8kcal/mole K, and for Klenow is ~ -0.9 kcal/mole K. Circular dichroism (CD) spectra for DNA binding by Taq/Klentaq show small but observable

conformational rearrangements of both the DNA and the protein upon binding. For Klenow, CD and SAXS spectra also show structural rearrangements of the protein and/or the DNA however, the changes are more subtle than for Klentaq. The structural data show a clear induced fit or coupled binding and folding event associated with DNA binding by both the polymerases, but one that is too small to account for all the ΔCp using the ΔCp and ΔASA correlations (1-4).

Some of the thermodynamic similarities between the two polymerases are as striking as any of their differences. For example, high negative ΔCps are usually considered to be a 'signature' of sequence specific DNA binding due to the highly complementary nature of the protein-DNA interface (5), yet the polymerases being predominantly non-sequence specific, bind DNA with high negative ΔCp. Sequence specificity of Klenow for the single-stranded template overhang was examined by altering the sequence to poly-A, poly-T and poly-C respectively. Binding titrations with the altered sequences show the affinity difference from weakest (poly-A) to tightest (poly-T) is only about one order of magnitude. In contrast, preliminary studies of the binding of Klenow and Klentaq to different DNA structures show that the two polymerases have very different DNA structure preferences (see Appendix 1 of this dissertation). We thus propose that the negative ΔCp observed in this study is due to DNA structure specificity, although it is also possible that it might be a general characteristic of primarily non sequence specific DNA binding proteins that bind with high affinity. The negative ΔCp observed for the DNA binding by Taq/Klentaq also exhibits temperature dependence. We propose that the restricted vibrational modes of the

protein upon DNA binding might contribute to the temperature dependence of this thermodynamic parameter.

The thermodynamic number of ions released upon DNA binding by these polymerases are relatively low compared to other DNA binding proteins. For example, the non-specific binding of lac repressor is associated with the release of ~9-11 monovalent ions (6), and the site specific binding to the operator is associated with the release of ~8 monovalent ions. It has been shown that these are predominantly cation releases, originating from the polyelectrolyte effect (6). Our [KCl] linkage analysis reveals a net release of 2-3 ions upon DNA binding of Taq/Klentaq, and a net release 4-5 ions upon DNA binding of Klenow. Linkage analysis using [KAc^-] shows a similar ion release, indicating that there is no anion specific effect (at least for these two anions) (Greg Thompson, unpublished data). Although the number of ions released are low compared to the number of bases covered upon DNA binding by the polymerases, if we consider the number of positively charged residues (Arg and Lys) whose side chains are within 5Å of the DNA phosphates, based on the co-crystal structures (PDB ID, 1KLN (for Klenow) and PDB ID 4KTQ (for Klentaq)) (7,8), then for Klenow we find that about 5 positively charged residues (Arg631, Lys635, Arg455, Lys422, Arg835) are involved (7,9), while for Klentaq 7 (Arg746, Arg728, Arg753, Lys540, Arg536, Arg487, and Lys508) are involved (8). For Klenow, out of these 5, a maximum of two residues could be involved in salt-bridge disruption with an adjacent acidic residue upon DNA binding. For Klentaq there are a maximum of 5 possible salt-bridge disruptions upon DNA binding (See Figure 7.1). These would result in the uptake of cations, by the acidic residues in the protein, therefore lowering the ion linkage (10). Even though there are

more positive charges within 5Å of the DNA in Klentaq, there are also more potentially disrupted salt bridges. Thus, the lower ion release for Klentaq could simply be due to the lower net value of the difference between the number of phosphate-side chain interactions and the number of disrupted salt bridges. Such a proposal, however, must be viewed with extreme caution in this situation, because the Klenow-DNA co-crystal is an editing complex, hence the comparison is not between equivalent bound complexes. (The positive residues in contact with the DNA are obtained from Ref. 8 and 9 and the other residues within 5Å of these residues are obtained by using the CSU (Contact of Structural Units) software in the RCSB website (11).

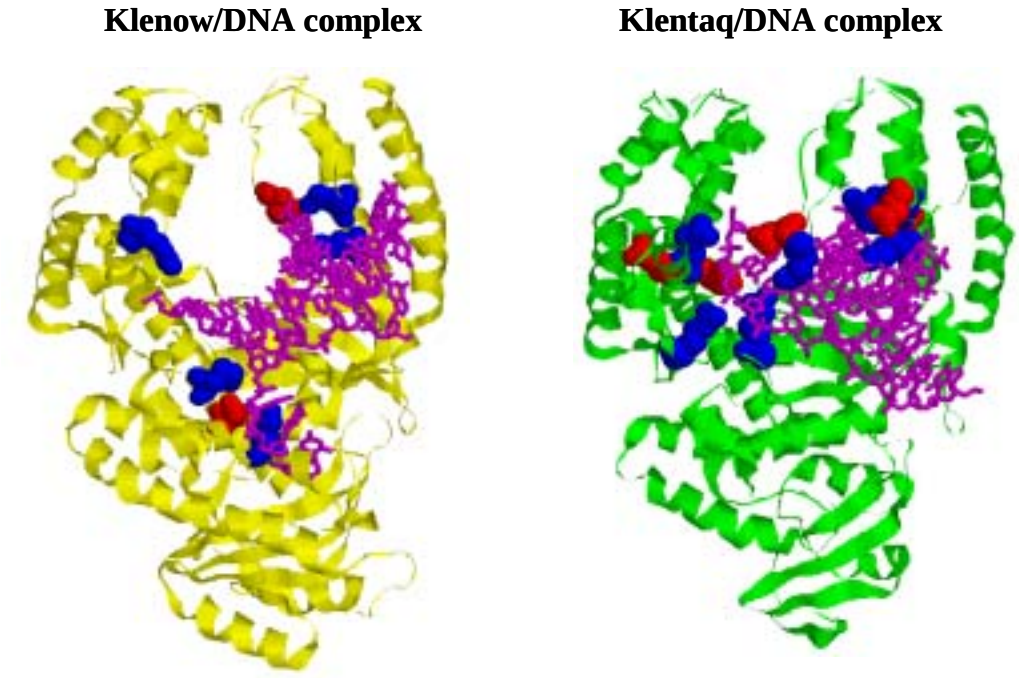


Figure 7.1. The DNA bound structures of Klenow (PDB ID 1KLN) and Klentaq (PDB ID 4KTQ), showing the positively charged residues (in blue) involved in DNA binding. The residues in red are the acidic residues that can form salt-bridges with the positive residues.

Although our experiments do not allow us to determine the exact DNA binding mode employed by Klenow in the conditions used in our studies (since Klenow can bind in both polymerization and editing modes), the dramatic differences in the DNA binding affinities of the two polymerases under similar solution conditions and the differences in the ions released, led us to suggest that the initial DNA recognition sites might be different for the two species of polymerases. Based on these results we speculated that the two polymerases might bind DNA with different footprints, a hypothesis we are currently testing by performing DNA footprinting experiments under similar salt and solution conditions for the two polymerases and also as a function of different solution conditions.

The DNA binding affinity also depends on the on-rate (k_{on}) and off-rate (k_{off}), for DNA binding, $K_d = k_{off}/k_{on.}$ Benkovic and coworkers have shown that for the 13/20mer DNA, the k_{on} for DNA binding by Klenow is very fast and is in the range of $1.2x10^7 M^{-1} s^{-1}$ while the off-rate is significantly slower, $\sim 0.06 s^{-1}$ (12). Studying the salt dependence of the k_{off} and k_{on} for the two polymerases will also potentially help us to understand the origin of the observed affinity differences for DNA binding by the two polymerases.

Our studies on the temperature dependence of DNA binding by the two species of polymerases have shown that their DNA binding thermodynamics are very similar to sequence-specific DNA binding proteins. We proposed that this might be due to the structure-specific binding, analogous to sequence-specific binding, and that this seems to be a characteristic of DNA polymerases in general. DNA structure specificity had been observed earlier for activity of the 5'nuclease domain of Taq polymerase (13,14). It was proposed that structure-specific binding of the 5'nuclease to damaged DNA might help to

minimize the indiscriminate cleavage of DNA by the polymerase (13,14). Although these studies did not assay for direct binding of the 5'nuclease domain, they clearly indicate that DNA structural specificity seems to be important for the proper functioning of the polymerases. Thus, DNA binding of the 5'nuclease domain might be expected to exhibit site-specific thermodynamics, as observed in our study.

Our preliminary data also suggests that the two polymerases differ substantially in their preferences for the DNA structures. We find that Taq/Klentaq has very similar preferences for primed-template and double-stranded DNA while Klenow has a significantly stronger preference for binding to primed-template DNA as compared to double-stranded DNA (see Appendix 1). This further suggests that the initial DNA recognition sites for the two polymerases differ. Further in-depth characterization of the DNA binding thermodynamics of these two species of polymerases with different DNA structures will not only help us to better understand the generalities and specificities of structure-specific DNA binding but may also provide with some plausible explanation for the differences in the DNA structure preferences by the two polymerases.

References

1. Spolar, R.S., Livingstone, J.R. and Record, M.T. (1992) *Biochemistry* **31**, 3947-3955.

2. Murphy, K.P. and Freire, E. (1992) *Adv. Prot. Chem*. **43**, 313-361.

3. Makhatadze, G.I. and Privalov, P.L. (1995) *Adv. Prot. Chem*. **47**, 307-425.

4. Myers, J. K., Pace, C. N. and Scholtz, J. M. (1995) *Protein Sci*. **4**, 2138-2148.

5. Jen-Jacobsen, L., Engler, L.E., Ames, J. T., Kurpiewski, M.R. and Grigorescu, A. (2000) *Supramolecular Chemistry* **12**, 143-160.

6. Ha, J.H., Capp, M.W., Hohenwalter, M.D., Baskerville, M. and Record, M.T. Jr. (1992) *J. Mol. Biol*. **228**, 252-264.

7. Beese, L.S., Derbyshire, V. and Steitz, T.A. (1993) *Science* **260**, 352-355.

8. Li, Y., Korolev, S. and Waksman, G. (1998) *EMBO J.* **17,** 7514-7525.

9. Eom, S.H., Wang, J. and Steitz, T.A. (1996) *Nature* **382,** 278-281.

10. Holbrook, J.A., Tsodikov, O.V., Saeker, R.M. and Record, M.T. (2001) *J. Mol. Biol*. **310**, 379-401.

11. Sobolev, V., Sorokine, A., Prilusky J., Abola, E.E. and Edelman, M. (1999) *Bioinformatics*, **15**, 327-332.

12. Kuchta, R.D., Mizrahi, V., Benkovic, P.A., Johnson, K.A. and Benkovic, S.J. (1987) *Biochemistry*. **26**, 8410-8417.

13. Lyamichev, V., Brow, M.A.D. and Dahlberg, J.E. (1993) *Science* **260**, 778-783.

14. Kaiser, M.W., Lyamicheva, N., Ma, W., Miller, C., Bruce, N., Fors, L. and Lyamichev, V. (1999) *J. Biol. Chem*. **274,** 21387-21394.

APPENDIX 1

PRELIMINARY DATA

In this appendix, I have described studies that are still in progress or in development in the ongoing characterization of the differences and similarities between DNA binding by type1 DNA polymerases from *Thermus aquaticus* and *Escherichia coli.* The general purpose of this appendix is to document the current status of some of these as yet incomplete studies to serve as background information and starting points for possible future development of these different investigations.

DNA Structure Preferences for Binding of Taq and Klenow DNA Polymerases

Introduction

Our previous studies have suggested that DNA structural preference might be correlated with the specific thermodynamic profiles observed for the binding of Taq/Klentaq and Klenow DNA polymerases to primed-template DNA (Chapters 4 and 5 of this dissertation). In order to better understand the relation between this proposed structure-specific DNA binding and the respective thermodynamic profiles, we have undertaken a study of the binding of the polymerases to different DNA constructs. The parallel studies with Taq/Klentaq and Klenow polymerases will also allow us to determine if there are any differential preference of the two polymerases for binding to specific DNA structures and how that might relate to the functioning of the proteins in their respective physiological conditions. Here we have examined the binding of Taq/Klentaq and Klenow to primed-template and double-stranded DNA. Our preliminary results show that the two species of polymerases differ significantly in their preferences for binding these two DNA structures.

Materials and Methods

The different DNA constructs used are the 13/20mer and 63/70mer DNA primed-template DNA (see Chapter 3) and the double-stranded 20/20mer. The 20/20mer DNA was constructed by adding the complementary bases to the primer strand of the 13/20mer DNA (Fig A.1). All the DNA constructs used for this study were unlabeled. The binding of the polymerases to different DNA constructs was examined using the electrophoretic gel-mobility shift assay under stoichiometric conditions. The proteins were added to unlabeled DNA in the binding buffer (10mM Tris, 125mM KCl, 5mM $MgCl_2$, pH 7.9) containing100ug/mL BSA. The reaction mixture was incubated at room temperature for 2 min and loaded onto a 4-20% TBE gel (BioRad), which was pre-run in 1X TBE buffer (89mM Tris Borate, 2mM EDTA, pH8.3) + 1mM $MgCl_2$ at 4^{o}C. A final concentration of 6% glycerol was added to the reaction mixture for loading onto the gel. No dye was added to the samples. The gel was run at a constant voltage of 100 V at 4^{o}C. The gel was then stained with Sybr Green (Molecular Probes) for 30mins and the image was obtained in a Nucleotech gel imager. The calorimetric enthalpies of binding of Klenow polymerase to the different DNA constructs were measured using isothermal titration calorimetry (described in Chapter 3) over a temperature range of 10 to 35^{o}C under identical buffer conditions (10mM Tris, 5mM $MgCl_2$, 300mM KCl, pH 7.9).

Results and Discussion

Lanes 3 and 4 of the gel-mobility shift assay in Figure A.1 show that Klenow binds to 13/20mer but does not form an electrophoretically stable complex with the double-stranded 20/20mer DNA. Previous studies have also shown that the DNA binding

affinities for Klenow and T4 DNA polymerases are dramatically reduced for ds-DNA as compared to primed-template DNA (1,2). In contrast, Taq/Klentaq can bind efficiently to

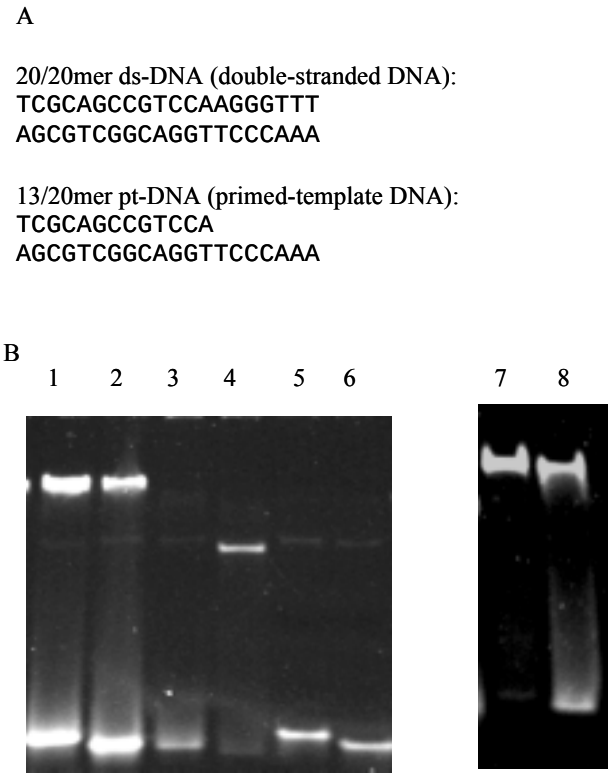


Figure A.1. A), The DNA constructs used for the assay. B), Gel shift binding assay showing the binding to the different DNA constructs. Upper bands are protein-DNA complexes and the lower bands are the free DNA. Lane 1: Taq + ds-DNA, lane 2: Taq + pt-DNA, lane 3: Klenow + ds DNA, lane 4: Klenow + pt-DNA, lane 5: Only ds-DNA, lane 6: Only pt-DNA, lane 7: Klentaq + ds-DNA, lane 8: Klentaq + pt-DNA.

both 13/20mer and 20/20mer DNA (Fig A.1, lanes 1 and 2). Equilibrium DNA binding studies using the fluorescence anisotropy assay show that the binding affinities of Klentaq for ds- and primed template DNA are almost identical (A.Wowor, data not shown). The calorimetric enthalpies (ΔH_{cal}) for binding of Klenow to the different DNA constructs as a function of temperature are shown in Fig. A.2. At all the temperatures

examined, the enthalpies of binding are closer to zero for the ds-DNA yielding a less negative ΔCp as compared to the primed-template DNA (See Table A.1 below).

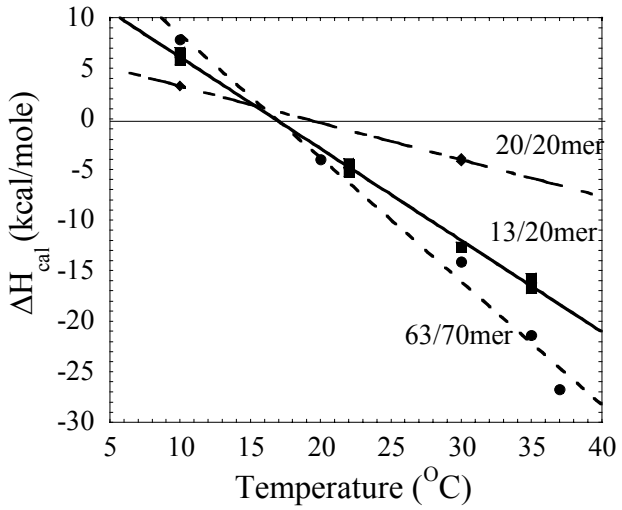


Figure A.2. Temperature dependencies of the calorimetric enthalpy changes (ΔH_{cal}) upon binding of Klenow to 20/20mer (♦), 13/20mer (■) and 63/70mer (●) DNA. The slopes of these plots yield the heat capacity changes (ΔCps) associated with the binding and the values are listed in Table A.1.

Table A.1. Calorimetric ΔH and ΔCp Values for DNA Binding by Klenow

13/20mer

Temperature(°C)	ΔH_{cal} (kcal/mole)	ΔCp_{cal} (kcal/mol K)
	+5.82	
10	+6.60	
22	-4.41	-0.91 ± 0.02
22	-5.25	
30	-12.66	
35	-15.73	
35	-16.74	

(table cont'd)

63/70mer

Temperature(^{o}C)	ΔH_{cal} (kcal/mole)	ΔCp_{cal} (kcal/mol K)
10	+7.85	
20	-3.99	-1.22 ± 0.08
30	-14.13	
35	-21.38	
37	-26.72	

20/20mer

Temperature(^{o}C)	ΔH_{cal} (kcal/mole)	ΔCp_{cal} (kcal/mol K)
10	+3.25	
30	-3.91	-0.36 ± 0.01
30	-4.16	

Our preliminary data on the structural specificity of DNA binding by Taq/Klentaq and Klenow polymerases indicate that the initial DNA structure recognition by the two polymerases are significantly different. Taq/Klentaq binds to ds- and primed template DNA with similar preferences and comparable thermodynamic profiles, suggesting that this polymerase may bind primarily to the double-stranded region of a DNA construct. In contrast, the reduced binding of Klenow to ds-DNA as compared to primed-template DNA indicates that DNA binding by this polymerase is directed either by the primer-template junction or by the single-stranded overhang of the template DNA. Previous studies have shown that ~4-5 bases of the template overhang are required for the optimal DNA binding affinity by Klenow and T4 DNA polymerases (1,2), further suggesting that

the DNA structure determinant for binding of Klenow differs from Taq/Klentaq. Although it is not clear as yet whether the ds-DNA binding by Taq/Klentaq is a general characteristic of non-proofreading DNA polymerases or a more specific property of Taq, the differences in DNA structural preferences between Taq and Klenow however suggest that the initial DNA recognition mechanism for the two polymerases are significantly different.

It is interesting to note that the respective similarities and differences in the DNA structural preferences by the Taq/Klentaq and Klenow polymerases, observed in this study, are also reflected in the ΔCp of binding to the corresponding DNA structures. For example, the binding of Klentaq to both ds- and primed-template DNA is associated with large negative ΔCps (data not shown), while for Klenow, the ΔCp of binding to ds-DNA is less negative as compared to primed-template DNA (Fig A.2). Solution based structural studies like DNA footprinting and SAXS experiments, of the binding of these polymerases to different DNA constructs, are currently in progress in order to determine the possible structural origins of the observed negative ΔCps of binding.

Osmotic Stress Experiments: Methods and Preliminary Data

Introduction

The specificity and stability of protein-DNA interactions are strongly influenced by the solution environment. Thus, studying the dependencies of equilibrium DNA binding as a function of various solution conditions provides us with important thermodynamic information for understanding the molecular mechanisms involved in the process. We have already characterized the salt and temperature dependencies of DNA binding by Taq and *E.coli* DNA polymerases (Chapter 3-5). The observed

thermodynamic profiles for binding of these polymerases to primed-template DNA (Chapter 4 and 5) and other preliminary data strongly suggest that these polymerases show DNA structural preference for binding. In this study we have examined the change in the macromolecular hydration accompanying the binding of Klentaq to primed-template DNA. The measured water stoichiometry will provide an estimate of the total changes in the solvent-accessible surface areas (ΔASA) of the reactants upon binding.

Crystal structures of many specific protein-DNA complexes show that direct protein-DNA contacts mostly replace the contacts of the free macromolecules with water with little or no water left at the interface (3-5). In contrast, the binding of sequence-specific proteins to a "non-cognate" sequence shows additional waters incorporated at the interface between the protein and the DNA (3,6). This is consistent with the observed reduction in the number of waters released for non-sequence-specific DNA binding as compared to specific binding (7 and 8-11). The DNA polymerases examined in this dissertation cannot be classified into any of the above two categories since they bind DNA primarily non-sequence-specifically but presumably with DNA structural preference (Chapter 4 and 5). Thus, understanding the participation of water in the binding of Klentaq to DNA will not only add to our present understanding of the DNA binding energetics of this polymerase, but would also be one of the first extensions of understanding the role of water in structure-specific DNA binding.

Here we have used the osmotic stress technique (12) in order to quantitate the change in the number of thermodynamic water molecules associated with DNA binding by Klentaq. This technique uses the thermodynamic linkage between osmotic pressure and the equilibrium of a macromolecular process to calculate the change in the number of

waters associated with that process (12-14). When the osmotic pressure of the bulk solution is increased in a controlled manner by using neutral solutes (osmolytes), movement of water away from "osmotically stressed" regions of the macromolecules are favored, so that a macromolecular process involving water release to the bulk solution is favored while one involving water uptake is inhibited (15). However, there are a few caveats that need to be considered while applying this technique for measuring the associated hydration change for any molecular process. The added solute only "stresses" those regions of the molecules from which it is excluded, such as the first hydration layer, surface crevices and internal cavities of the macromolecules. Thus, the solute size dependence of the osmotic stress effects has to be considered and solute sizes that will "stress" all the functionally significant spaces in the protein and the DNA should be used (See Ref. 12). Also, the binding equilibrium should not be perturbed by the preferential interaction of the solute with either the protein or the DNA. Another important caveat is that the volume exclusion by the osmolytes should not significantly alter the binding affinity (excluded volume effect) (12). Thus, upon eliminating these potential added contributions of the stressing solutes to the protein-DNA interaction (by performing well designed controls), the dependence of the affinity of the protein for the DNA on the osmotic pressure of the solution will give the measure of the net change of the number of water molecules that are involved in the process. We find that the binding of Klentaq to 13/20mer primed-template DNA is associated with a net release of ~587 waters. This number is ~1.6X larger than the number of waters predicted from the ΔASA calculations based on the x-ray crystal structures, according to the relation that each water molecule occupies ~7-10 $Å^2$ (14) of the macromolecular surfaces. It notable that the observed

water release is significantly higher than observed for non-specific protein-DNA interactions and is more comparable to sequence-specific DNA binding (7-11), suggesting that like sequence-specific DNA binding, the binding of Klentaq to DNA also involves tight packing of the interacting surfaces by excluding water from the interface (3). This is consistent with our previous finding showing that the binding of Klentaq to primed-template DNA is associated with a high negative ΔCp, a thermodynamic parameter believed to be correlated to the burial of extensive surface area upon protein-DNA interaction, also a signature of sequence-specific DNA binding (Chapter 4). Thus it seems that the basic DNA binding thermodynamics for structure-specific DNA binding (as proposed for Klentaq) and sequence-specific binding are very similar.

Materials and Methods

Materials

The Klentaq large fragment of Taq DNA polymerase was purified as described in Chapter 2 of this dissertation. The fluorescently labeled and unlabeled DNA constructs were purchased from Integrated DNA Technologies Inc. Polyethylene glycols, Dextrans and Ficoll 70 were purchased from Fluka Chemika-Biochemika (Milwaukee, WI).

Fluorescence Anisotropy Assay

Equilibrium DNA binding experiments were performed by titrating Klentaq into 1nM ROX-labeled 13/20mer DNA as described in Chapter 3, in a buffer containing 10mM Tris, 5mM $MgCl_2$, 100mM KCl, pH 7.9 (base buffer) at 25^{o}C with differing amounts of stressing solutes to produce the desired osmotic pressures. Both the DNA solution in the cuvette and the protein titrant solution contained osmolytes at the same

concentration. Proper care was taken to ensure thorough mixing of the solution inside the cuvette during titration.

Preparation of Buffers with the Osmolytes

For each of the osmolytes, stock solutions containing 30% (w/v) of each of the osmolytes in the base buffer was prepared. These were then diluted with the base buffer to produce the desired osmotic pressure. The pH of the buffers were carefully monitored and the addition of the solutes caused little or no change in pH. Osmotic pressures of the buffer solutions were measured on a Wescor vapor pressure osmometer. The osmotic pressure of the base buffer is 5.2 (±0.2) atm, which thus is the lowest osmotic pressure (effective origin) for all experiments.

Data Analysis

The equilibrium binding curves obtained using fluorescence anisotropy were fit to a standard single-site isotherm as described in Chapter 4.

Osmotic stress results were analyzed using thermodynamic relationships developed previously (12,14,16). Briefly, the change in the number of associated water molecules is obtained from the slope of the plot of $\ln(K^{\pi}/K^{0})$ versus osmotic pressure using the relationship developed previously (12,14,16):

$$\{\partial kT \ln(K^{\pi}/K^{0})\}/\{(\partial \pi_{osm})\} = \Delta V_w = \Delta N_w (30\text{Å}^3) \qquad \textbf{Eq. A.1}$$

where K^{π} is the association constant ($1/K_d$) at osmotic pressure π, K^0 is the association constant in the base buffer in the absence of any solute, ΔV_w is the linked change in volume, 30Å^3 is the molecular volume of one molecule of water, ΔN_w is the linked change in the number of associated waters.

Results

The stressing solutes used to determine the change in the associated water molecules for DNA binding by Klentaq are polyethylene glycols (PEGs, from M_w 300 to 8000), dextrans (from M_w 15,00 to 20,000) and Ficoll 70 (M_w 70,000). Since the added solute stresses only regions of the protein and the DNA from which they are excluded, it is important to first determine the size of the solute that would "stress" all the functionally significant hydration spaces in the protein and the DNA. This is achieved by examining the DNA binding affinity of the polymerase as a function of solute sizes at a particular osmotic pressure. Figure A.3 shows the PEG and Dextran size dependencies of DNA binding at an osmotic pressure of 7.8(±0.1) atm.

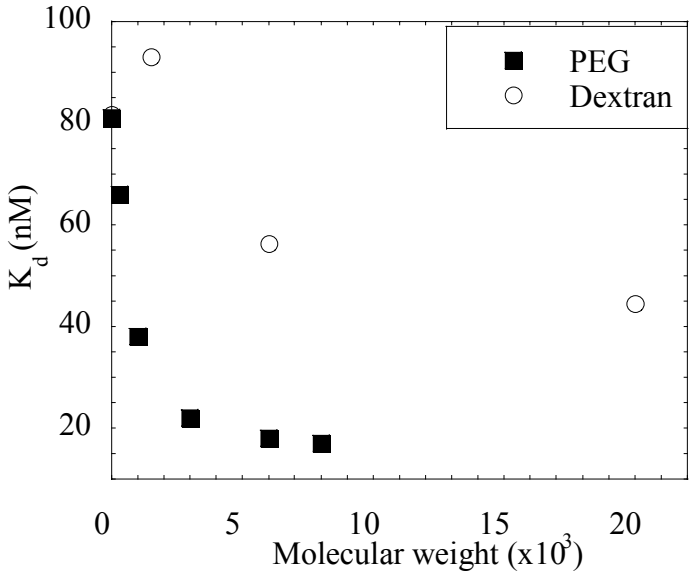


Figure A.3. Osmolyte size dependencies at constant osmotic pressure for DNA binding by Klentaq. All data on this plot were collected at 7.8(±0.1) atm osmotic pressure as induced by different molecular weights of PEG (■) and Dextran (○).

The binding affinities for different solute sizes are listed in Table A.2. The binding exhibits significant size dependency with both PEG and dextran. For both the osmolytes, the binding gets tighter with increasing solute size up until a certain point after which it levels off. For PEG the effect on the DNA binding plateaus at PEG size of ≥ 6000, however a higher solute size had to be used for dextran to eliminate the solute size effect on DNA binding. The plateauing of the solute size dependence indicates that, 1) any solute size in the plateau region would be equally effective in probing the full hydration changes associated with the DNA binding and 2) for any solute size in the plateau, the molecular crowding effect would not be the primary effect of added solute (as this effect would continue to increase with added solute size) (reviewed in detail in Ref.15). This led us to use PEG 6000 in order to determine the osmotic stress effect on the DNA binding by Klentaq for estimating the water linkage associated with the process.

We also investigated the effects of two other stressing solutes, Ficoll 70 and Dextran 6000, on the binding affinity. Figure A.4 shows the osmotic pressure dependencies of the DNA binding affinity in PEG 6000, Ficoll 70 and Dextran 6000 respectively. For all the solutes examined, the binding affinity gets tighter with increasing osmotic pressure indicating a net release of water upon DNA binding. Analyses of these linkage plots using Eq. A.1 yield the number of water released and the values are reported in Table A.3. Similar estimates of water release are reported for both PEG 6000 and Ficoll 70. The similar water linkages observed for PEG 6000 (a synthetic polymer of ethylene glycol) and Ficoll 70 (a neutral, highly branched hydrophilic polymer of sucrose) which are chemically significantly different, indicate the absence of any significant solute specific effects, and suggest the values reported from these solutes are

likely the true estimates of the water release. The solute size effects in Ficoll were not examined because Ficoll-70 is considered to be large enough to probe the overall hydration changes in macromolecular processes and is extensively used for this purpose by Bloomfield and coworkers (17).

The water release reported from Dextran 6000 in our study is significantly lower than for Ficoll 70 or PEG 6000. Figure A.3 shows that the osmotic stress effects in Dextran solutions has not yet plateaued for Dextrans of 6000Mw. Thus Dextran 6000 is not an appropriate stressing solute for assaying the total water release for this reaction. The difficulties of working with high concentrations of high molecular weight Dextran

Table A.2. **Solute (Osmolyte) Size Effects on the Binding of Klentaq to 13/20mer Primed Template DNA**

All the titrations with the osmolytes were performed at a constant osmotic pressure of 7.8(±0.1) atm.

Nature of solute	Concentration of solute % (w/v)	K_d (nM)
No osmolyte	0	81.7 ± 1.2
PEG 300	2	66.2 ± 1.6
PEG 1000	6	38.7 ± 0.4
PEG 3000	9	22.4± 0.8
PEG 6000	10	18.3 ± 0.4
PEG 8000	10	17.9 ± 0.4
Dextran 1500	8	93.1 ± 2.8
Dextran 6000	15	56.3 ± 0.9
Dextran 15-20,000	20	44.5 ± 1.1

solutions, restricted the use of Dextran 15-20,000 as a stressing osmolyte for studying the osmotic pressure effects on the binding of Klentaq.

The results with Dextran 6000, however, are a striking example of how important the establishment of the solute size effect plateau is for each reaction and solute combination examined using osmotic stress. Examples exist where use of solute sizes smaller than those on the empirically determined size effect plateau can even result in apparent movement of water in the complete opposite direction (18).

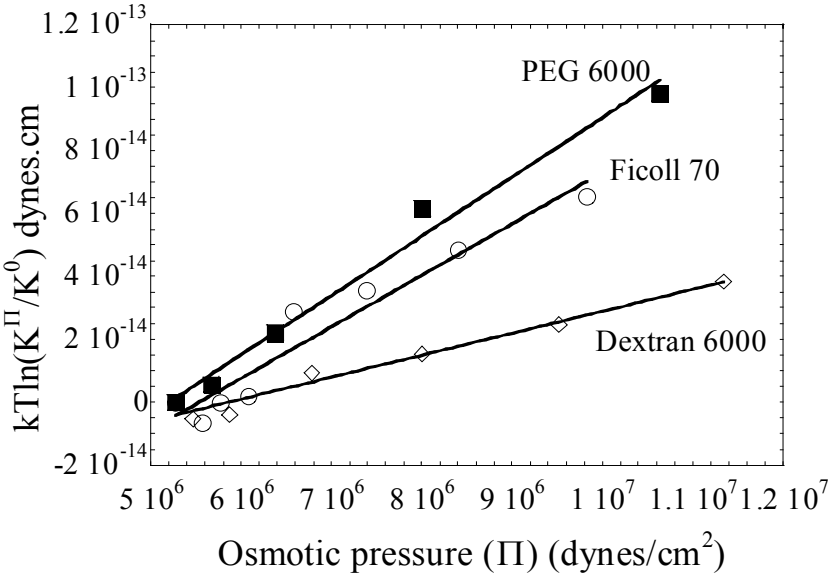


Figure A.4. Osmotic pressure dependencies of the binding of Klentaq to 13/20mer DNA, plotted as described in Eq.A.1 in the text. The solute used to raise the osmotic pressure are: PEG 6000 (■), Ficoll 70 (○) and Dextran 6000 (◊). The osmotic pressure of the buffer without any added solute is 5.2(±0.2) atm.

Discussion

Typically, protein-DNA interactions are associated with net release of water molecules (however, see Ref 36 for an exception), primarily due to the displacement of water from the protein-DNA interfaces. The release of water also increases the entropy of

Table A.3. Change in Associated Water (ΔH_2O) Measured by Osmotic Stress[a]

Nature of solute	Concentration of solute % (wt/v)	π (atm)	K_d (nM)	ΔH_2O [a]
No osmolyte	0	5.2 (± 0.2)	81.7 ± 1.2	
PEG 6000	2	5.6	71.2 ± 1.6	
	5	6.3	47.6 ± 0.8	628 ± 47
	10	7.9	18.3 ± 0.4	
	15	10.5	7.5 ± 0.2	
Ficoll 70	2	5.5	95.8 ± 3.2	
	5	5.7	81.8 ± 2.7	
	10	6.0	78 ± 2.2	
	15	6.5	40.6 ± 1.1	546 ± 58
	20	7.3	34.4 ± 1.6	
	25	8.3	25.1 ± 1.3	
	30	9.7	16.6 ± 0.9	
Dextran 6000	2	5.4	93.2 ± 2.9	
	5	5.8	89.8 ± 1.4	
	10	6.7	65.2 ± 2.1	233 ± 19
	15	7.9	56.3 ± 0.9	
	20	9.4	44.7 ± 1.6	
	25	11.2	32.1 ± 0.9	

[a]ΔH_2O and associated errors are calculated by fitting to Eq.A.1 in the text.

the system, thereby contributing favorably to the binding free energy. The osmotic stress technique, developed by Parsegian and coworkers (12,13), allows the determination of the hydration changes associated with any macromolecular process. Our present study of

the osmotic pressure dependence for the binding of Klentaq to primed-template DNA predicts a net release of ~587 water molecules (mean for the values obtained for PEG 6000 and Ficoll 70). Although it is not entirely clear how the water linkage estimated from the osmotic stress measurements scales with other molecular parameters, it is most often correlated with the solvent-accessible surface area change (ΔASA) associated with the process. Based on the results obtained with proteins, it is estimated that each water molecule occupies ~7-10 $Å^2$ of macromolecular surface (14). From this relation, the observed release of ~587 water molecules would imply a ΔASA of ~5000 $Å^2$. This number is larger than the ΔASA calculated from the x-ray crystal structures that predicted a value of ~3216 $Å^2$ (see Chapter 4 of this dissertation). Although, the accurate calculation of surface area and volume changes based on x-ray crystal structures is an area of active investigation, we can postulate some origins for this observed discrepancy. 1) The estimate that one water molecule occupies ~7-10 $Å^2$ of macromolecular surface is based on results obtained from studies with proteins. It has been proposed previously that the area occupied by a water molecule on a highly charged DNA surface may be smaller due to electroconstriction (20). Thus, as has been pointed out in the earlier studies on lac-repressor-DNA interaction (7), the estimation of the ΔASA from osmotic stress measurements, using only protein-derived parameters, is most likely an upper limit for the total ΔASA. 2) The observed ΔASA for a protein-DNA interaction typically reflects the burial of surfaces at the interface plus any conformational changes in the protein and /or the DNA that buries additional surface areas. The structure-based ΔASA value reported above accounts for the interface area and the conformational changes in the

protein but does not include the contributions from the observed DNA structural rearrangements.

Although this study does not allow us to clearly define the exact molecular origins of the observed discrepancy, it is interesting to note that the total water linkage observed for Klentaq-DNA interaction is very large, again a general characteristic of sequence-specific DNA binding. This is consistent with our previous results where, again analogous to sequence-specific binding, the DNA binding of Klentaq is associated with a large negative ΔCp (Chapter 4). Since Klentaq binds DNA primarily non-specifically, the present finding further corroborates our previous hypothesis proposing that, in analogy to sequence-specific binding, the DNA binding thermodynamics observed for Klentaq are correlated with structure-specific binding.

Reference

1. Turner, R.M., Grindley, N.D.F. and Joyce, C.M. (2003) *Biochemistry* **42**, 2373-2385.

2. Delagoutte, E. and von Hippel, P.H. (2003) *J. Biol. Chem*. **278**, 25435-25447.

3. Sidorova, N.Y. and Rau, D.C. (1996) *Proc. Natl. Acad. Sci. USA*. **93**, 12272-12277.

4. Steitz, T.A. (1990) *Q. Rev. Biophys*. 23, 205-280.

5. Pabo, C.O. and Sauer, R.T. (1992) *Annu. Rev. Biochem*. **61**, 1053-1095.

6. Gewirth, D.T. and Sigler, P.B. (1995) *Nat. Struct. Biol*. **2**, 386-394.

7. Fried, M.G., Stickle, D.F., Smirnakis, K.V., Adams, C., MacDonald, D. and Ponzy, L. (2000) *J. Biol. Chem*. **277**, 50676-50682.

8. Garner, M.M. and Rau, D.C. (1995) *EMBO J*. **14**, 1257-1263.

9. Vossen, K.M., Wolz, R., Daugherty, M.A. and Fried, M.G. (1997) *Biochemistry* **36**, 11640-11647.

10. Robinson, C.R. and Sligar, S.G. (1995) *Methods Enzymol*. **259**, 395-426.

11. Brown, M.P., Grillo, A.D. and Royer, C.A. (1999) *Protein Sci*. **8**, 1276-1285.

12. Parsegian, V.A., Rand, R.P. and Rau, D.C. (1995) *Methods Enzymol*. **259**, 43-94.

13. Parsegian, V.A., Rand, R.P., Fuller, N.L. and Rau, D.C. (1986) *Methods Enzymol*. **127**, 400-416.

14. Colombo, M.F., Rau, D.C. and Parsegian, V.A. (1992) *Science* **256**, 655-659.

15. LiCata, V.J. and Allewell, N.A. (1997) *Biochemistry* **36**, 10161-10167.

16. Rand, R.P., Fuller, N.L., Butko, P., Francis, G. and Nicholls, P. (1993) *Biochemistry* **32**, 5925.

17. Bloomfield, V.A. and Wenner, J.R. (1999) *Biophys. J*. **77**, 3234-3241.

18. LiCata, V.J. and Allewell, N.M. (1998) *Methods Enzymol*. **295**, 42-62.

19. Robinson C.R. and Sligar, S.G. (1996) *Protein Sci*. **5**, 2119-2124.

20. Gerstein, M., Tsai, J. and Levitt, M. (1995) *J. Mol. Biol*. **249**, 959-966.

APPENDIX 2

COPYRIGHT RELEASE PERMISSION

Copyright Release for Chapter 3

Department of Biological Sciences
Louisiana State University
107 Life Sciences Building
Baton Rouge, LA 70803
Phone: (225) 5781589
Fax: (225) 5782597

February 2, 2004

Charles C. Hancock
Executive Officer
American Society for Biochemistry
and Molecular Biology
9650 Rockville Pike
Bethesda, MD 20814-3996

Dear Sir:

This letter will confirm our recent email communication (see attachment). I am completing a doctoral dissertation at Louisiana State University entitled "Thermodynamic characterization of DNA binding by Type 1 DNA Polymerases from *Thermus aquaticus* and *Escherichia coli*." I would like your permission to reprint in my dissertation excerpts from the following:

> Kausiki Datta and Vince J. LiCata. Salt Dependence of DNA binding by *Thermus aquaticus* and *Escherichia coli* DNA Polymerases. *The Journal of Biological Chemistry*. **Vol. 278**, No: 8, Issue of February 21, 2003. pp. 5694-5701.
>
> The excerpts to be reproduced are: The full text of the paper including the figures and the tables.

The requested permission extends to any future revisions and editions of my dissertation, including non-exclusive world rights in all languages, and to the prospective publication of my dissertation by UMI Company. These rights will in no way restrict republication of the material in any other form by you or by others authorized by you. Your signing of this letter will also confirm that your company owns the copyright to the above-described material.

If these arrangements meet with your approval, please sign this letter where indicated below and return it to me in the enclosed return envelope. Thank you very much.

Sincerely,
Kausiki Datta
Kausiki Datta

PERMISSION GRANTED FOR THE USE REQUESTED ABOVE:

American Society for Biochemistry
and Molecular Biology

By: _____

Title: _____

Date: FEB 09 2003

PERMISSION GRANTED
E. Oppenheimer
for the copyright owner
THE AMERICAN SOCIETY FOR BIOCHEMISTRY
& MOLECULAR BIOLOGY

Copyright Release for Chapter 4

OXFORD
UNIVERSITY PRESS

Rights and New Business Development, Journals
Great Clarendon Street, Oxford OX2 6DP, UK
Telephone +44 |0| 1865 354490 ***new***
Fax: +44 |0| 1865 353485 ***new***
Email chris.payne@oupjournals.org
Ref: SMJ/FP/NARESE/Datta/CP/0104

12/02/04

Kausiki Datta
Dept. of Biological Sciences
Louisiana State University
107 Life Sciences Bldg.
Baton Rouge, LA 70803
USA

Dear Kausiki Datta

RE: Nucleic Acids Research, Vol. 31 (19), 2003, pp. 5590-5597
Datta & LiCata, 'Thermodynamics of the binding…'

Thank you for your letter dated 2- February 2004, requesting permission to reprint the above material. Our permission is granted without fee to reproduce the material and to distribute the article via the UMI, as you are the original author.

Use of the **article** is restricted to the title 'Thermodynamic characterization of DNA…', available in *print & microform* format only, to be used only in the *English* Language. This permission is limited to this particular use and does not allow you to use it elsewhere or in any other format other than specified above.

Please include a credit line in your publication citing full details of the Oxford University Press publication which is the source of the material and by permission of Oxford University Press or the sponsoring society if this is a society journal.

If the credit line or acknowledgement in our publication indicates that material including any illustrations/figures etc was drawn or modified from an earlier source it will be necessary for you to also clear permission with the original publisher. If this permission has not been obtained, please note that this material cannot be included in your publication/photocopies.

Please do not hesitate to contact me if I can be of any further assistance.

Yours sincerely,

Chris Payne
Journals Rights Assistant
(on behalf of Fiona Willis, Journals Rights and New Business Development Manager)

For an online permissions forms visit our web site at:
http://www3.oup.co.uk/jnls/permissions/

VITA

Kausiki Datta was born on February 1, 1974, in Calcutta, India. Kausiki received her bachelor of science in 1995 and the master degree in 1997 from the University of Calcutta, India. She worked as a project assistant in the division of protein engineering in Indian Institute of Chemical Biology, Calcutta, from 1997 to 1999. In 1999 she entered the graduate program in Louisiana State University in the Department of Biological Sciences as a doctoral student. As a graduate student Kausiki worked in the laboratory of Dr. Vince J. LiCata towards her dissertation project and also taught introductory biology courses in the department. After graduation in May 2004, Kausiki plans to join the laboratory of Dr. Peter H. von Hippel in the University of Oregon as a post-doctoral research fellow.